META-ANALYSIS OF ENVIRONMENTAL VALUATION STUDIES

A Dissertation
Presented to
The Academic Faculty

by

Sheldon Gen

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in Public Policy

Georgia Institute of Technology
December 2004

META-ANALYSIS OF ENVIRONMENTAL VALUATION STUDIES

Approved by

Dr. Bryan Norton, Chair

Dr. Michael Farmer

Dr. William Bommer

Dr. Robert Kirkman

Dr. Michael Elliott

November 19, 2004

ACKNOWLEDGEMENTS

I am indebted to several people who have provided guidance and support for this research, and I wish to express my deep thanks to them. Most importantly, I thank my teachers at Georgia Tech – especially Bryan Norton and Gordon Kingsley – and my dissertation committee for inspiring me to ask questions and guiding me to seek answers. I also thank my cohorts, who have taught me as much as any class. I especially thank Jim Dietz, Elizabeth Corley, Jongwon Park, Kathy Brice, Kamau Bobb, and Leora Waldner. Finally, I am thankful for the financial support for this research, provided in part by the U.S. Environmental Protection Agency's National Network for Environmental Management Studies and the Georgia Chapter of the Air & Waste Management Association.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

α	level of significance
χ^2	chi square, statistic of the χ^2 test of independence
δ	population effect size, derivative
μ	population mean
AJAE	American Journal of Agricultural Economics
ANOVA	analysis of variance
b	unstandardized regression coefficient
CA	California
CBA	cost-benefit analysis
C.I.	confidence interval
CO	Colorado
CV	contingent valuation
d	sample effect size
D	average effect size
DC	District of Columbia
d.f.	degrees of freedom
e.g.	for example
ERE	Environmental and Resource Economics
et al.	and others
etc.	etcetera
F	statistic of the analysis of variance test

H_0	null hypothesis
H_1	alternative hypothesis
HP	hedonic pricing
i.e.	that is
IL	Illinois
JEM	Journal of Environmental Management
K	number of effect sizes examined
LE	Land Economics
LSD	least significant difference
ME	Maine
MA	Massachusetts
n	sample size
N	average sample size
NA	not applicable
NEBA	net environmental benefit analysis
NOAA	National Oceanic and Atmospheric Administration
OR	Oregon
p.	page
PA	Pennsylvania
pp.	pages
Q	statistic of the χ^2 test of heterogeneity
r	Pearson's correlation coefficient, implicit price in hedonic pricing
R^2	coefficient of multiple determination

RUM	random utility model
S	standard deviation
t	statistic of the difference in means test
T	total sample size
TC	travel cost
U.S.	United States
Var(d)	weighted variance of d
Var(e)	sampling error variance
WC	weak complementarity
WI	Wisconsin
WRR	Water Resources Research
WTA	willingness to accept
WTP	willingness to pay
WY	Wyoming
X	independent variable
Y	dependent variable
$Y_E - Y_C$	change in value due to environmental improvement

SUMMARY

Cost-benefit analysis has long been a popular method of policy analysis. When applied to policies affecting the environment, however, it faced a serious problem. Many policies exert environmental impacts that are external to market transactions, so the monetary values of these impacts resist measurement. Specifically, price and quantity information, normally the demand and associated values of normal goods, are missing for many environmental goods. This shortcoming in cost-benefit analysis has led some researchers to develop indirect and nonmarket methods of valuing environmental goods monetarily. Three prevalent methods have been developed and widely used over the last few decades: the travel cost method, hedonic pricing, and contingent valuation. However, these tools' development and growth in popularity over the last few decades have been accompanied by much debate over the reliability and validity of the methods and their results. This dissertation research tests the convergent validity of these popular environmental valuation tools through a meta-analysis of effect sizes from a sample of studies using these tools. The results bring statistical evidence to the debate over their uses, and refine the conditions of their implementation by identifying factors that affect their results.

Tools of Environmental Valuation

The travel cost method recognizes that travel costs are an important component of the full cost of visiting any site and enjoying certain environmental goods, and that there is wide variation in the travel costs for visitors. The method thus indirectly values

environmental goods by analyzing consumption behavior in markets directly related to the good. The costs of obtaining or enjoying the services of the environmental asset are used as a proxy for price. The method is conducted through surveys of travelers to sites with environmental amenities to be valued. The surveys ask travelers questions about their trips to the sites, the purposes of the trips, the distances traveled, the time and money spent, and their socio-economic characteristics. From the survey data, researchers estimate demand curves for the sites and extrapolate their values to relevant populations (Freeman, 1993, p. 445). The method assumes weak complementarity between the environmental good and the consumption expenditure. It is thus possible to infer the value of a change in the environmental good from the demand for the market good.

A major advantage of this method of environmental valuation is that it is based on real, observable market behavior, giving the method some degree of construct validity. A key problem with the travel cost method, however, is that it assumes that users of the environmental goods do perceive and respond to the variations in the levels and qualities of them (David, 1971; Freeman, 1993, p. 447). Also, weak complementarity methods are properly interpreted as only the lower bounds of maximum willingness to pay.

Hedonic pricing is also a weak complementarity method. It recognizes that sometimes the quantity or quality of an environmental good can be an attribute of a market good. In these cases, information on the value of the environmental good is embedded in the prices and consumption levels for the marketed goods, and analysis of this information can extract the inferred value of the environmental good (Freeman, 1993, p. 124). The method is conducted through statistical analysis of the marketed good, which is usually housing. Prices for homes are modeled as functions of several characteristics including environmental amenities. Through regression analysis of the data and differential calculus the method isolates the relationships between the levels of environmental goods and the prices of the homes.

Like the travel cost method, a major advantage of hedonic pricing is that it indirectly measures the values of nonmarket goods from real, observable market behavior (Desvousges, Johnson, & Banzhaf, 1998, p. 17). However, also like the travel cost method, one of its major disadvantages is its reliance on the assumption that consumers of the related market goods perceive the differences in environmental amenities, which might not be evident (David, 1971; Hite, 1998). Like other weak complementarity methods, hedonic pricing also underestimates maximum willingness to pay.

The third method, contingent valuation, takes a very straightforward approach to valuing environmental goods. It directly asks people the maximum amount they would pay for an improvement in an environmental amenity, or the minimum amount of payment they would accept for a reduction in an environmental amenity, if it were traded in a market. The responses to the question then represent the value of the environmental change (Freeman, 1993, p. 165). In this method, values are not indirectly inferred from market behavior, but directly asked outside of a real market. Data for contingent valuation are gathered through carefully structured surveys that describe hypothetical markets in which the goods are traded. After responses to contingent valuation surveys are gathered, the data can then be analyzed to estimate mean values, test hypotheses about the influence of income and other demographic characteristics on value, and estimate total benefits through extrapolation of results to the relevant populations.

Contingent valuation is unlike the hedonic pricing and travel cost methods in very important ways. First, it is able to measure nonuse values, values people place on environmental goods from which they gain no actual use. Use values, on the other hand, are those derived from the current direct or indirect use of the amenity. In contrast to the contingent valuation method, the hedonic pricing and travel cost methods can only measure use values. Second, data for contingent valuation come from *stated preferences* of respondents, rather than from market behaviors. In contrast, the hedonic pricing and travel cost methods are based on *revealed preferences* from the subjects' real market consumption of weak complements of the environmental goods. Contingent valuation's reliance on stated preferences makes the method susceptible to strategic responses. Finally, economic theory predicts that welfare estimates from contingent valuation should be different than those from hedonic pricing and the travel cost method, but there is dispute over the direction of the difference.

Popularity and Controversy

All three of these methods have facilitated the use of cost-benefit analysis in environmental decision making, and they have been used extensively in recent decades. This is most obviously evident in the numbers of environmental valuation studies performed. For example, Carson, Wright, Alberini, Carson, and Flores (1994) compiled an extensive bibliography on just one method of environmental valuation, contingent valuation, and found 1,672 articles. Overall, these environmental valuation efforts have been used for cost-benefit analyses of government actions, environmental damage assessment, environmental costing, and environmental accounting (Navrud & Pruckner, 1997).

The use of these tools has also been popularized and highlighted in benefit transfer studies. These are policy analyses that use results of existing valuation studies, implemented for their own specific purposes and contexts, to evaluate policy choices in other contexts (Brookshire & Neill, 1992). The most obvious advantages of the method are its economies of money and time (Desvousges, Johnson, & Banzhaf, 1998, pp. 9-10; Freeman, 1993, p. 484; Ofiara & Seneca, 2001, p. 264). Indeed, the benefit transfer method has become a "bedrock" (Desvousges, Johnson, & Banzhaf, 1998, p. 1) of environmental policy analysis, and the valuation studies supporting the method have become useful beyond their original intent.

Despite the popularity of these environmental valuation methods and their use in environmental cost-benefit analysis, they are controversial. Economists and others have taken all sides of the discussion on the reliability and validity of the methods and their normative implications. On one extreme, strong proponents of environmental valuation and cost-benefit analysis place much confidence in the theory behind the tools used. For Freeman (1993) and other leaders of environmental valuation, state of the art tools such as the travel cost method, hedonic pricing, and contingent valuation represent valid approaches to infer monetary value of environmental goods. Furthermore, utilizing these tools allows government decision makers to effectively address a wide range of environmental policy problems such as setting efficient levels of environmental standards and analyzing the net benefits of existing and proposed regulations. Less enthusiastic than Freeman but still supportive of environmental valuation are proponents who are concerned with the vulnerabilities of the tools but concede the necessity of measuring the monetary values of environmental goods in a market economy. The National Oceanic and Atmospheric Administration's blue ribbon panel on contingent valuation exemplify this position. The panel endorsed contingent valuation as a means of valuing environmental goods, but they tempered their endorsement with a long list of procedural cautions stemming from the technical vulnerabilities of the method (Arrow et al., 1993).

On the other side of the debate are a range of opponents to cost-benefit analysis applied to environmental policies and environmental valuation itself. Marginal opponents are those who like the idea of including monetary valuations of the environment in policy analyses, but see too many problems with the methods currently available (e.g., Green & Tunstall, 1991; Kellert, 1984; Neill, Cummings, Ganderton, Harrison, & McGuckin, 1994). They note deficiencies in the environmental valuation tools that proponents of the tools also recognize. Some highlight the cognitive difficulties of answering valuation questions in contingent valuation (Desvousges, Johnson, & Banzhaf, 1998, pp. 15-16). Others point to the inability of revealed preference methods to include nonuse values in their measures, as well as the unverifiable and unsatisfactory state of the art of the one method that can measure it (Goodland & Ledec, 1994, p. 452). Perhaps the most common criticisms of current environmental valuation tools are the inconsistencies and biases in their results (e.g., Balistreri, McClelland, Poe, & Schulze, 2001; Boyle, MacDonald, Cheng, & McCollum, 1998; Kealy, Dovidio, & Rockel, 1988; Loomis, Brown, Lucero, & Peterson, 1996; Loomis, Brown, Lucero, & Peterson, 1997; Seip & Strand, 1992). King (1998) argues that the results of environmental valuation efforts can be so unreliable and subjective that their use in policy decision making has been detrimental to the environment instead of helpful.

Finally, at the other extreme of the continuum, there are strong opponents of environmental valuation who object to the very idea of monetizing environmental values. Their arguments are mostly normative, pointing to the inappropriateness of utilitarian ethics applied to environmental policies. Kelman (1998), for example, argues that environmental goods are priceless and not for sale, so cost-benefit analysis is not applicable. Instead, he claims, environmental policies are more appropriately judged by deontological criteria. Sagoff (1997, 1981) concurs, arguing that the true value people place on environmental goods is not confined to market prices nor to people's desires as consumers. Instead, when people judge the value of environmental goods they do so from their obligations as citizens.

As if the above controversy over the validity and ethics of environmental valuation were not enough, the debate's fire has been recently fanned by the growth of benefit transfer studies. The method assumes the generalizability of the existing literature (Smith, 1992), but does not necessarily verify it. This is problematic because most applications of the travel cost, hedonic pricing, and contingent valuation methods have been used to evaluate specific, local environmental changes, with their generalizability only qualitatively assessed or left to the judgment of their readers altogether.

Given the popularity of these methods and the controversy surrounding them, it is surprising that there have been very few meta-analyses of valuation studies. Existing meta-analyses have tended to focus on specific goods or on specific valuation tools.

However, these studies attempted to find the *determinants* of value, rather than test the validity of the methods. Indeed, Smith and Pattanayak (2002) reviewed 15 meta-analyses of nonmarket valuation and found that 13 of them sought to synthesize valuations of specific goods and identify the determinants of the values.

One meta-analysis, however, attempted to address the validity of two valuation methods (Carson, Flores, Martin, & Wright, 1996). Its findings generally support the validity of the methods, but they were based only on studies that compared stated and revealed preference methods together. No one has yet compared the results of *different* stated preference and weak complementarity studies for the purpose of testing the convergent validity of the valuation tools. Furthermore, none of the existing meta-analyses have used effect sizes as the outcome measure of valuation studies, a common practice in meta-analyses that facilitates a broader examination of original studies. This dissertation research fills the gap in the meta-analyses of environmental economics by testing the convergent validity of three environmental valuation methods and refining their proper contextual uses.

Method

The guiding questions of this research are, "are the results from different environmental valuation methods collectively reliable, are there contexts in which they are more reliable than others, and – to the extent possible – are the results valid?" Thus, there are two objectives for this research. First, it tests the convergent validity of the valuation methods. Convergence of results is not a sufficient determinant of validity, but it is a necessary condition. While measuring convergent validity, this research also identifies moderating variables explaining the variance in the valuations. Second, this research folds the results of these statistical analyses into the broader normative discussion surrounding environmental valuation.

To accomplish these objectives, I performed an extensive meta-analysis of environmental valuation studies. The statistical mechanics of meta-analysis varies with the information available from the studies evaluated, but the framework of all meta-analyses is common (Cooper & Hedges, 1994, pp. 9-13; Hunt, 1997; Hunter & Schmidt, 1990). It consists of collecting studies applicable to the research question, coding information from them, and analyzing the coded data. The primary hypotheses tested by this method are the existence of moderating variables explaining the variance among the studies' effect sizes. The primary variable of interest is the method of valuation. This essentially is the test of convergent validity, because if all three methods validly measure environmental values, then the method used should not moderate value. Other variables tested for their moderating effects include the type of environmental good, the magnitude of environmental change, the description of environmental change, the years of the data, and the locations of the studies.

The above hypotheses were tested by statistical analysis of data collected from the existing body of environmental valuation studies. Following the meta-analytic framework, the research began with a thorough and systematic search for applicable studies, followed by the coding of information from the studies. Studies included in the data set met three criteria: they were published in leading environmental valuation journals, employed at least one of the three valuation methods of interest, and quantitatively valued changes in environmental goods.

I searched five bibliographic databases pertaining to environmental issues. In each database I systematically searched for journal articles pertaining to the travel cost method, hedonic pricing, and contingent valuation. The journals cited in these searches were recorded, as well as the number of hits for each journal. Altogether, 141 journals were cited, with 753 hits (including multiple hits from more than one database). However, over a third of all hits came from just five journals. These top five journals were manually searched. To capture the most current practices used in each of the valuation methods, I started the literature search from December 2001 and worked back to the beginnings of the contemporary practices of each method, which provided a sufficient sample for the statistical analyses. From each study meeting the criteria of inclusion a list of variables was coded, and these variables provided the data for the statistical analyses.

Analysis of the resulting data consisted of descriptive statistics, calculations of effect sizes, and tests for moderating variables. Effect sizes are standardized measures of the impact the independent variable has on the dependent variable. In this research the effect sizes are differences in valuation for an environmental good that is subject to degradation or improvement. Calculations of effect sizes vary according to the information reported in the original studies, but follow the general form

$$d = (Y_E - Y_C)/S,$$

where Y_E - Y_C is the difference in average effect (valuation) before and after the treatment (change in environmental quality), and S is the pooled standard deviation from both cases (Hunter & Schmidt, 1990, pp. 233-235). The effect size is unitless, thus facilitating comparability between studies. With effect sizes calculated, the heart of the meta-analysis proceeded with the combining of effect sizes while adjusting for sample size, and testing for the moderating effect of variables

Data and Results

The manual search resulted in over 400 articles dealing with some aspect of these environmental valuation methods. Further examination of these resulted in a total of 228 articles that met all the selection criteria. These 228 articles were analyzed in depth and their data were coded. The 228 cases resulted in a total of 614 valuations of changes in environmental goods. With these data, the body of environmental valuation studies were characterized as such:

- Water, land, and air are well represented in the valuation studies. Animals are too, but the diversity of species valued make their comparability, and the transferability of their results, questionable. Other environmental goods such as plants, toxics, and wastes, have received relatively little attention in environmental valuation.

- Contingent valuation is the most popularly used valuation method, representing a majority of the valuations in this data set.

- Different methods tend to be used in different situations. Specifically, contingent valuation tends to be used with qualitative descriptions of holistic changes in environmental goods. While it is used to value all types of goods, it is especially favored when valuing animals. Hedonic pricing tends to be used with quantitative descriptions of incremental changes. Its use has been limited to valuing water, land, and especially air. The travel cost method is rarely used to value changes in environmental goods.

- Certain journals disproportionately publish valuations from different places, of different goods, or using different methods.

- Although environmental valuation outside of the U.S. has received little attention, the attention it has received is heavily focused on land values, and less so on air or animal values. Also, valuations outside the U.S. are more likely to employ contingent valuation than hedonic pricing or the travel cost method.

- There has been a sharp rise in environmental valuation studies since the mid-1980s. The rise has been characterized by studies on land and water goods, and studies employing contingent valuation. The rise coincides with a few events – especially the maturation of contingent valuation – that have helped make environmental valuation popular.

For each observation with sufficient reported information an effect size was calculated. The effect sizes were corrected for sampling error and combined, following Hunter and Schmidt's (1990) procedures. The average effect size weighted by sample size is 6.90. This number indicates that, based on this data set, the average change in value of an environmental good due to its improvement is 6.90 standard deviations. This summary statement overlooks all the variance there is in the studies generating this data set, but the direction and order of magnitude of the mean effect size is instructive. It tells us that, assuming the validity of the valuation methods (an assumption questioned in this research), people do place a significant monetary value on improved conditions of the environment. However, tests of homogeneity strongly suggested the presence of moderating variables. Thus, the hypothesized moderators were next tested, including environmental good, magnitude of change, description of change, valuation method, year of data collection, and location of study. Those found to be significant were included in regression analyses of effect sizes, to determine the relative effects of the moderators. These analyses resulted in the following conclusions:

- Type of good is a significant moderator. Air and water have significantly larger average effect sizes than land and animal, and air has a significantly larger average effect size than water.

- Magnitude of change and type of description of change are significant moderators, but their effects are mixed. These conflicting results may be due to validity problems in the measures themselves.

- The method of valuation, the primary variable in this research, is a significant moderator. Effect sizes from contingent valuations are, on average, significantly smaller than those from the travel cost and hedonic pricing methods. There is no significant difference between the effect sizes of the travel cost and hedonic pricing methods. On average, contingent valuation produces effect sizes that average 40% to 55% less than those of the hedonic pricing and travel cost methods. The difference varies with the environmental goods valued. There is evidence, however, that contingent valuations have become more reliable over time. Even so, the difference between the results of contingent valuations and travel cost or hedonic pricing methods remain significant. There are competing explanations for this difference, including the researchers' designs, the respondents' replies, and the scopes of the methods' measurements.

- The year of data collection is not a significant modifier. There is not significant evidence to demonstrate changes in environmental values over the years.

- Location of study is a significant moderator. Effect sizes in the United States average twice the magnitude of all other countries in the data set, but the relatively small number of foreign studies in the data set casts doubt on this finding. Within the United States effect sizes in California are, on average, significantly different than those from other states in the data set. Specifically, effect sizes for air tend to be much smaller in California, and those for land tend to be much larger.

While these results are supported by much empirical evidence, they do have limitations. The most apparent are the unexplained variance in the regression models and the differing subsets of studies used for different analyses, which present possible spurious effects in the analyses. A common limitation of any meta-analysis is the possibility of publication bias, but this is not of great concern to this meta-analysis because of the size of the data set and because several studies did report insignificant effects. A final limitation of these results is perhaps the most important: its generalizability. The sample of studies in this research is not representative, but were selected from the most prolific journals in environmental valuation to presumably capture high quality studies. This strategy was used to qualitatively minimize the chances of a type 1 error. The consequence of this strategy is that the magnitude of the divergence is not representative. The advantage, however, is that the conclusion of divergence is better substantiated.

Implications

Following these analyses, the most fundamental question that must now be asked is, "What do environmental valuations tell us?" The evidence in this research does not completely support one side of the debate over environmental valuation, but it certainly leans with the critics. Interestingly, the evidence brings attention to a second question that appears equally important: "*Whose* values do they represent?"

Addressing the first question, proponents and opponents agree that environmental valuations are intended to measure economic values of environmental goods. Total economic value of environmental goods is composed of direct use values, indirect use values, option values, existence values, and bequest values (Barbier, 1994). But the three methods of valuation do not all claim to measure total economic value, only portions of it. The hedonic pricing and travel cost methods can only measure direct use values.

Contingent valuation, in contrast, is claimed to be able to measure all types of economic value, but these distinctions are often absent in its applications and left to the judgments of its readers. Thus, the first conclusion that can be made is that environmental valuations speak of direct use economic values, and in some cases (contingent valuation) it may also speak of other dimensions of economic value.

How authoritatively do they speak of these values? The evidence from this study is mixed. It clearly says that environmental values are positive. The consistency with which positive measures of values were attributed to environmental improvements lends some credibility to the methods. That is, the methods measure the *direction* of value well[1]. But what can be said about the *magnitudes*? Do they accurately express the same economic values of environmental goods? The divergence of effect sizes among the methods say no. Contingent valuations and the two weak complementarity methods produce significantly different average effect sizes, while controlling for environmental good, holistic versus incremental changes, and quantitative versus qualitative measures of change.

At first glance, the answer to the second question seems simple: the values expressed or revealed are those of the subjects, be they respondents in a contingent valuation study, home buyers in a hedonic pricing study, or visitors in a travel cost study. But this interpretation belies the contextual problem of the valuation methods. The results of contingent valuations are significantly influenced by the information provided by the researchers. A similar problem exists for the hedonic pricing and travel cost methods. The selections of goods to include in the travel cost and hedonic pricing models are, in essence, researchers' judgments of the weak complementary bundles of goods consumers are purchasing with their properties and travels. The reality of what consumers are consciously purchasing may range from none of the environmental goods identified by the researchers to entire spectra of economic values of the goods. For all three methods, researchers frame the environmental goods being valued by identifying the goods, specifying their scopes, and selecting the measures of them. None of these tasks is an objective decision, and all of them influence or define the results of valuation studies. Thus, when assessing the meanings of the outputs of environmental valuations, it is important to recognize that the values expressed are not only those of the subjects, but also those of the researchers.

One value of researchers that is consistently expressed in the applications of these methods is in the scope of value itself. For each of these three methods, value is limited to an economic dimension and is measured in monetary terms. However, there is significant evidence that people object to expressing their values for environmental goods in solely economic terms, and find it difficult to do so (Baron & Spranca, 1997; Kahneman & Knetsch, 1992; Sagoff, 1998; Stevens, Echeverria, Glass, Hager, & More, 1991). Thus, it seems clear that the values of environmental valuation researchers can be in conflict with those of the people they study.

Given these mixed results, what uses of these valuation methods remain defensible? Others have described guidelines for the proper application of environmental valuation tools, and the results from this study provide greater detail to these guidelines. Considering Navrud and Pruckner's (1997) hierarchy of need for accuracy in

---

[1] This claim is based on the assumption that values for environmental improvements are indeed positive, which is substantiated by opinion surveys on environmental values (e.g.

environmental valuation and Norton's (1995) environmental risk decision square, along with evidence from this research, the defensible uses of the environmental valuation methods are quite narrow. They can be appropriate in cases of local reversible impacts, and in such cases, hedonic pricing and travel cost methods are more reliable choices when accuracy is demanded. But, again, these two methods only capture direct use values and ignore all nonuse values. When accuracy is not needed, then estimates of total economic value from contingent valuation may stimulate discussions for local, reversible impacts. But when *both* accuracy and total economic value are needed, none of these three methods will suffice.

The results of this research also speak to the defensible uses of existing valuation studies in benefit transfers. The literature already identifies key study qualities to consider when transferring results from one context to another. They include the environmental good, the change in the good, the time in which the values were measured, the location of the study, the quality of the study, the demographics of the beneficiaries and the population studied, and the hypothetical markets in contingent valuations. All of these factors are presumed to moderate value, and the analyses in this research qualify and quantify the effects of some of them. But most importantly, this research adds the method of valuation to the list of factors to consider when transferring benefits. Overall, contingent valuations produced effect sizes that averaged 40% to 55% less than those of the hedonic pricing and travel cost methods. However, the effect varies substantially by environmental good. Ultimately, benefit transfers are vulnerable to the same validity issues as the original studies, but to a greater degree because of the additional issue of the

---

Kempton, Boster, & Hartley, 1999).

comparability of the circumstances in the original studies and the target situation. Thus benefit transfers should be limited to situations demanding low accuracy and relative economic values.

What environmental valuations aim to tell us is a limited, economic view of value. However, economic values are not as relevant to environment policy as the popularity of environmental valuations suggests. Instead, the environment elicits many other dimensions of value that have been vying for a substantial voice in public policy. To the extent that environmental valuations limit the discussions to only economic considerations, they are detrimental to environmental decision making. Instead, valuations of the environment ought to facilitate expressions of many dimensions of value. The literature again provides guidance on other values relevant to environmental decision making. They include communal and public values to complement individuals' private values (Bozeman, 2002; Gowdy, 1997; Norton, Costanza, & Bishop, 1998; Sagoff, 1997; Stone, 1997), cooperative values to complement competitive values (McLaughlin, 2003; Shabman & Stephenson, 1996), deontological values to complement utilitarian values (Milbrath, 2003), and environmental quality measures to complement personal wealth measures (Brady, 1983; Brower, 1995). Taken together, the relevant measures of environmental values include economic terms, non-economic social terms, and environmental terms. They represent a pluralistic approach to replace one that has been dominated by economic considerations (Gowdy, 1997; Norgaard, 1985). These prescriptions for what environmental valuations ought to tell us begs the question, "How can these components of value be measured?" While finding a single measure to capture all aspects of relevant values is unlikely and perhaps undesirable, alternative methods are available to capture the relevant values with a combination of measures.

Good methods of measuring environmental values are those that facilitate public decision making by 1) articulating those values, 2) making tradeoffs among conflicting goals, 3) understanding sources of conflict, 4) solving distributional problems, 5) integrating values with technical analysis, and 6) anticipating future consequences (Maguire, 2002). By themselves, the three methods of economic valuation in this research do ostensibly well in facilitating tradeoffs, but they fail miserably on all other counts. However, a pluralistic approach to environmental valuation that incorporates some of the alternative methods introduced here can meet all of these criteria, and thus improve environmental policy decision making processes.

CHAPTER 1

INTRODUCTION

On the northern coast of Alaska, bordering the Beaufort Sea, is the remote Arctic Coastal Plain area of the Arctic National Wildlife Refuge. The 1.5 million acre area is a calving ground for the Porcupine River caribou and ecologists describe it as the greatest primeval wilderness ecosystem in the United States (Markey, 2001). Recognizing this unique ecology, neighboring Canada has protected the parts of the habitat lying on its side of the international border. However, geologists describe the Coastal Plain as the greatest oil and gas reserve in the United States. Just to the west of the Coastal Plain are the Prudhoe Bay oil fields and the beginning of the great Alaska Pipeline. The fields have been developed for oil and gas extraction for over 20 years and they currently constitute 21% of the nation's crude oil production (Young, 2001). Oil is transported from Prudhoe Bay south through the Alaska Pipeline, a major artery in the country's life-blood of oil.

Thus the Coastal Plain is the coveted prize in an ongoing legislative battle being fought in Congress. On one side are proponents of opening the Coastal Plain for oil and gas development. Representative Don Young of Alaska led these efforts on January 3, 2001. Following a campaign promise of President George W. Bush, Young introduced in the House of Representatives bill HR39, The Arctic Coastal Plain Domestic Energy Security Act. Among other things, the bill proposed to repeal prohibitions against oil and gas development in the Coastal Plain and order the Department of the Interior to lease the land for such purposes. Young's Alaskan colleague in the Senate at that time, four-term

Senator Frank Murkowski, introduced corresponding bill S388 in the upper house on February 26.

On the other side were proponents of preserving the Coastal Plain in its natural state. Anticipating the efforts to open the protected land to development, preservationists quickly introduced alternative bills in both houses. On February 28, 2001, Representative Ed Markey of Massachusetts introduced bill HR770, the Morris K. Udall Arctic Wilderness Act. If passed it would have designated the Coastal Plain area as wilderness for permanent preservation in its undeveloped state. Senator Joseph Lieberman of Connecticut introduced companion bill S411 in the Senate on the same day.

All four bills were under review in the houses' committees on environment and natural resources at the close of the 107th Congress. Three of them were reintroduced in the 108th Congress, but none left its committee for a floor vote.[1] With President Bush's reelection in November, 2004, it is likely that this battle will continue in the 109th Congress.

The debate over these alternatives has included assessments of the costs and benefits of each, and thus this battle for the Coastal Plain illustrates the problematic intersection of environmental policy and decision analysis. The benefits and costs of these two alternative policies are difficult to measure on comparable terms, especially in the measure of choice: money. Some of the outcomes of the alternatives are easy to quantify monetarily. Proponents of oil and gas development, for example, cite the expected savings from developing the Coastal Plain, as well as the costs of not developing it (Young, 2001). They claim the production of 11.6 to 31.5 billion barrels of

---

[1] SR388 in the 107th Congress was not reintroduced in the 108th Congress.

oil at prices below those of imported oil, and the reduction of the current $100 billion per year bill on foreign energy. However, other important aspects of the alternatives resist quantification. Proponents of preserving the Coastal Plains, for example, cite benefits such as the existence of unique wildlife and wilderness, and the preservation of the "wilderness heritage" to "bequeath undisturbed to future generations" (Markey, 2001, section 2.a.1).

Faced with the need to quantify the values of such environmental qualities, researchers developed a few tools to extrapolate their monetary values from behaviors and responses of consumers, and thus complete the calculations of costs and benefits of policies affecting the environment. However, these tools' development and growth in popularity over the last few decades have been accompanied by much debate over the reliability and validity of the methods and their results, as well as the appropriateness of monetizing environmental values. This research tests the convergent validity of three popular environmental valuation tools through a meta-analysis of effect sizes from a sample of studies using the valuation tools. The results of this research complement those of the few other meta-analyses of environmental valuation, by using effect sizes as the outcome measures of original studies, a common practice in other fields using meta-analysis. This approach offers a broader examination of environmental valuation studies that brings statistical evidence to the debate over their uses, and refines the conditions of their implementation by identifying factors that affect their results.

The remaining part of this chapter describes in detail the problem of cost-benefit analysis in environmental policy analysis, then introduces the three environmental valuation tools that are analyzed in this research. Chapter 2 describes the history and growth of these valuation tools, and discusses in detail the academic and public policy controversies that currently surround them. Chapter 3 describes the meta-analytic framework that is used to test these valuation tools. It discusses in detail the hypotheses that were tested, the method of collecting data to test the hypotheses, and the statistical analyses that were performed. Chapter 4 summarizes the data collected for this research and their descriptive statistics. Chapter 5 details the results of the hypotheses tested. Finally, Chapter 6 folds the results of the analyses back into the broader debate of valuation in environmental policy analysis, and discusses their implications on the use of the valuation tools.

Cost-Benefit Analysis and the Environment

Cost-benefit analysis has long been a popular method of policy analysis. In the United States, assessments of the costs and benefits of government projects began as early as 1808 (Hines, 1973). Formal prescription of cost-benefit analysis as a means of justifying public projects was first enacted in the Flood Control Act of 1936 (Ofiara & Seneca, 2001; Parsons, 1995). It required the benefits of any adopted action to exceed the costs. In recent history, formal cost-benefit analyses of all newly proposed regulations were prescribed in 1981 through President Ronald Reagan's Executive Order 12291. Today, cost-benefit analysis is the most popularly used tool of decision analysis (Munger, 2000, p. 352). Part of the popularity of this method of policy analysis resides in the ubiquity of its well understood measure – money – and its quantitative nature (Patton & Sawicki, 1993).

Simply put, cost-benefit analysis is an economic tool used to measure the monetary worthiness of actions. Its method follows the rational framework for policy analysis (Bardach, 2000; Patton & Sawicki, 1992; Tong, 1986) and fundamentally consists of calculating the benefits and costs of alternative actions, and then for each alternative comparing the two sums against each other (Desvousges, Johnson, & Banzhaf, 1998, p. 10; Munger, 2000, p. 356; Ofiara & Seneca, 2001, p. 41; Paris & Reynolds, 1983). The usual decision rule is to reject alternatives whose costs are greater than their benefits, and to accept actions whose benefits are greater than their costs. When more than one alternative have greater benefits than costs, then the one providing the greatest benefit (Paris & Reynolds, 1983; Parsons, 1995, p. 400), or the greatest benefit per unit cost, is most acceptable.

When applied to policies affecting the environment, however, cost-benefit analysis faced a serious problem. The root of the problem lies in the measure of costs and benefits. Standard texts in the method prescribe the measure of outcomes in monetary terms (Munger, 2000, p. 352; Weimer & Vining, 1992, p. 260), or at least recognize that it is standard practice to do so (Freeman, 1997, p. 189). But many costs and benefits in environmental policies are not easily calculated in monetary terms because the market does not usually include prices or payments directly for environmental goods. In his popular book *Earth in the Balance*, then-Senator Al Gore described the problem this way:

> "The hard truth is that our economic system is partially blind. It 'sees' some things and not others. It carefully measures and keeps track of the value of those things most important to buyers and sellers, such as food, clothing, manufactured goods, work, and, indeed, money itself. But its intricate calculations often completely ignore the value of other things that

> are harder to buy and sell: fresh water, clean air, the beauty of mountains, the rich diversity of life in the forest, just to name a few… just as our eyes fail to see all but a narrow portion of the light spectrum, our economics fail to see – let alone measure – the full value of major parts of our world." (Gore, 1992, p. 183)

Gore went on to exemplify the consequences of this problem of environmental economics:

> "In fact, the partial blindness of our current economic system is the single most powerful force behind what seem to be irrational decisions about the global environment… When we add up the costs and benefits of growing the grain, the loss of that freshwater resource will be ignored. And largely because we have failed to measure the value of clean, fresh groundwater, we have contaminated more than half of all the underground reservoirs in the United States with pesticide runoff and other poisonous residues that are virtually impossible to remove." (Gore, 1992, pp. 183-184)

In economic terms, this problem stems from environmental "externalities" of public policies and from the characteristics of "public goods" or "common property" often found in environmental goods and services. Many policies exert environmental impacts that are external to market transactions, so the monetary values of these impacts resist measurement (Desvousges, Johnson, & Banzhaf, 1998, p. 14; Freeman, 1993, pp. 2-3). Specifically, price and quantity information, normally the demand and associated values of normal goods, are missing for many environmental goods (Ofiara & Seneca, 2001, p. 197). For example, a traditional cost-benefit analysis of adopting technologies to reduce air pollution from industries might easily measure the costs of purchasing, installing, and maintaining such technology, but it would have much difficulty in measuring the monetary benefits of cleaner air because air is not directly traded in the market. This problem exists for all kinds of environmental goods not directly traded in the market, such as the quality of water (in rivers, lakes, aquifers, etc.), the health of ecosystems, the existence of endangered species, etc.

Furthermore, many environmental goods exhibit characteristics of public goods or at least common property: non-excludability and non-rivalry in consumption. For example, the provision of clean ambient air through tough pollution controls is non-excludable: once it is provided, its benefits cannot practically be limited to only paying consumers. Also, the clearer visibility from the clean air is non-rival in consumption: the "use" of the visibility by one consumer does not limit its availability for others. Many environmental goods also have the characteristic of not allowing individual consumers to specify their level of consumption. People cannot, for example, completely select the air quality around their homes or have much control over it.

All these characteristics of many environmental goods, not just clean air, make markets for them difficult to function properly (Freeman, 1993, p. 23) and can lead to their undervaluation and misuse as Gore (1992) described. In the example of the Coastal Plain of the Arctic National Wildlife Refuge, the environmental benefits of preserving the unique ecosystem, and the environmental costs of developing the area for oil and gas exploration, are mostly external to market transactions and exhibit public good characteristics. Accurate and precise measures of the policy alternatives' costs and benefits are thus elusive and controversial.

Sometimes in environmental policy we are interested in costs to the environment, such as the cases of environmental disasters that inflict harm. In these cases of natural resource damage assessments, the problem of monetary measurement also exists (Desvousges, Johnson, & Banzhaf, 1998, p. 10). A prime example of this lies at the other end of the Alaska Pipeline, away from Coastal Plain and at its southern terminus in Valdez, Alaska. On March 24, 1989, the oil tanker *Exxon Valdez* hit Bligh Reef in Prince William Sound, spilling 10 million gallons of oil into it. It was the largest oil spill in U.S. history, affecting 1,500 miles of coastline (US Senate Committee on Energy and Natural Resources, 1999, p. 2). In the aftermath of the spill, the federal and state governments sought to quantify the values of the economic and environmental damages from the spill. Monetary losses to fisheries, chartered boating, and other market services and goods were relatively easy to measure, but losses to the visual beauty and ecology of the area were much more difficult and suspect. Yet, measuring values of lost scenery and wildlife were of great importance because of the litigation that would result from the spill (Keeble, 1999, p. 197). Natural resource damage assessment is the process of quantifying and monetizing these damages to the environment (Ofiara & Seneca, 2001, p. 36), and it generally involves identifying the environmental resources services provided, measuring the reduction of these resources and services, and determining the values placed on them by those who had benefited from them (Freeman & Kopp, 1999, p. 52). However, environmental damage assessment suffers the same problem as environmental cost-benefit analysis: many environmental goods are not directly traded in the market, and they often exhibit characteristics of public goods.

Despite these difficulties in measuring the monetary values of environmental goods, the need to do so remains. Gore (1992, p. 346) called for changes in environmental policy analysis to include environmental costs and benefits, in order to fully account for the outcomes of our policies and to prevent undervalued uses of natural resources. As environmental economist A. Myrick Freeman III stated, “It is this failure of the market system to allocate and price resource and environmental services correctly that creates the need for economic measures of values to guide policymaking" (Freeman, 1993, pp. 2-3).

Three Tools of the Trade

This shortcoming in cost-benefit analysis has led some researchers to develop indirect and nonmarket methods of valuing environmental goods monetarily. Three prevalent methods have been developed and widely used over the last few decades: the travel cost method, hedonic pricing, and contingent valuation.

Travel Cost Method

The travel cost method recognizes that travel costs are an important component of the full cost of visiting any site or enjoying certain environmental goods, and that there is wide variation in the travel costs for visitors. The method thus indirectly values environmental goods by analyzing consumption behavior in markets directly related to the good. The costs of obtaining or enjoying the services of the environmental asset are used as a proxy for price.

The method is conducted through surveys of travelers to sites with environmental amenities to be valued. The surveys ask travelers questions about their trips to the sites, the purposes of the trips, the distances traveled, the time and money spent, and their socio-economic characteristics. From the survey data, researchers calculate travel costs

(C) as functions of distance costs (DC), time costs (TC), and entry fees (F) (Hanley & Spash, 1993, pp. 84-85):

$$C_{ij} = C(DC_{ij}, TC_{ij}, F_i), \qquad i = 1...n, j = 1...m$$

where “i” is each respondent and “j” is a given site. Researchers then estimate trip generation functions that predict the number of visits that will be taken by person “i” to site “j” as functions of the travel costs, socio-economic characteristics of the respondents, and the purpose of the trips. This multiple regression equation can then be used to estimate demand curves for the site and to extrapolate its value to the relevant population (Freeman, 1993, p. 445).

To make the link between the costs of enjoying an environmental amenity and the value of the amenity, the method assumes *weak complementarity* between the environmental good and the consumption expenditure. That is, the enjoyment of the environmental good requires the purchase of a market good. It is thus possible to infer the value of a change in the environmental good from the demand for the market good. These environmental and market goods are weak complements of each other (Freeman, 1993, p. 104). An example is traveling to enjoy a remote lake: enjoyment of the lake requires the purchase of a travel mode.

The travel cost method is the oldest of the three methods discussed. The method was conceived by Harold Hotelling in the late 1940s (Desvousges, Johnson, & Banzhaf, 1998, p. 20; Hanley & Spash, 1993, p. 83) but was not formally developed until the late 1950s through the mid-1960s (Freeman, 1993, p. 444; Hanley & Spash, 1993, p. 83). Since then the method has been most widely used in pricing outdoor recreation such as fishing, hiking, and camping (Desvousges, Johnson, & Banzhaf, 1998, p. 19), but it is also used to value broader environmental goods and changes in environmental qualities. It usually does the latter by one of two ways. One way recognizes that respondents have different levels of demand (i.e., numbers of visits) for similar sites with different levels or qualities of environmental amenities. Another way recognizes that respondents will change their demand for one site based on changes in the qualities of environmental amenities there. Either way, the responses of people to these variations become the basis for estimating the values of changes in the level or qualities of environmental goods (Freeman, 1993, pp. 443-444). This variation of the method to value changes in site quality is sometimes called the travel cost varying parameter model (Ofiara & Seneca, 2001, pp. 207-208).

For example, Choe, Whittington, and Lauria (1996) used the travel cost method to estimate value losses from degraded shore water quality in the Philippines. They surveyed residents near a beach recently contaminated with unsafe levels of fecal coliforms and pathogens. The survey asked respondents to report the number of visits to the beach they made per year before and after the contamination. The change in visitation rates, and their corresponding costs, represented the change in value. They found that the change in water quality from safe to unsafe levels of these pollutants resulted in a loss of visitations with a value of 36 to 51 pesos per person per month.

A major advantage of this method of environmental valuation is that it is based on real, observable market behavior, giving the method some degree of construct validity.

In contrast, one of the other two methods rely on responses to hypothetical markets[2] (Ofiara & Seneca, 2001). A key problem with the travel cost method, however, is that it assumes that users of the environmental goods do perceive and respond to the variations in the levels and qualities of them (David, 1971; Freeman, 1993, p. 447). Also, weak complementarity methods are biased, underestimates of maximum willingness to pay for environmental improvements. Their estimated environmental values are properly interpreted as lower bounds of willingness to pay. Other problems include the difficulty of including multipurpose trips, and how to value respondents' time (Hanley & Spash, 1993, pp. 86-91; McConnell, 1975).

Hedonic Pricing

Hedonic pricing is also a weak complementarity method. It recognizes that sometimes the quantity or quality of an environmental good can be an attribute of a market good. For example, a family can choose a vista with the house they buy. In these cases, consumers have some freedom to choose their consumption level of the environmental good through their selection of the marketed bundle. Information on the value of the environmental good is embedded in the prices and consumption levels for the marketed goods, and analysis of this information can extract the inferred value of the environmental good (Freeman, 1993, p. 124).

[2] Sometimes researchers add hypothetical market questions to a travel cost survey to gauge the change in behavior in response to some hypothetical future condition of the environmental good. For example, one might ask respondents how their number of trips to a site would change in response to some improvement to the site, like reduced air pollutants to increase visibility. However, such an approach – which some call the contingent activity method (Freeman, 1993, p. 166) – is a departure from the original method which is based on real market behavior.

The method is conducted through statistical analysis of the marketed good, which is usually housing. Prices for homes are modeled as functions of several characteristics including environmental amenities. Through regression analysis of the data and differential calculus the method isolates the relationships between the levels of environmental goods and the prices of the homes. Specifically, the prices of homes (P) are modeled as a function of three vectors of attributes: site characteristics (S) such as the numbers of rooms and lot size, neighborhood characteristics (N) such as crime rate and quality of schools, and environmental qualities (Q) such as waterfront and air quality (Hanley & Spash, 1993, p. 76):

$$P = P(S_i, N_j, Q_k), \qquad i = 1...m,\ j = 1...n,\ k = 1\text{-}l.$$

Next, this equation is partially differentiated with respect to Q to determine the implicit price, r, for the environmental good:

$$\delta P/\delta Q = r\ .$$

Finally, demand curves for Q can be estimated by regressing its implicit price on Q and relevant socio-economic characteristics (SE) such as age, income, and education:

$$r = P(Q_k, SE_l), \qquad k = 1...m,\ l = 1...\ n.$$

For example, Kiel (1995) used hedonic pricing to estimate the losses in land value from the presence of hazardous waste sites. To do so, she analyzed the Woburn, Massachusetts housing market and the two Superfund sites within that city. The prices of the houses were regressed on the minimum distances from each house to either

Superfund site. Interestingly, she repeated this analysis for six consecutive spans of time representing different stages of discovery and clean up of the hazardous waste sites. During the discovery phase, when odors from the sites were first noticed and wells in the area tested and closed, she found that a one mile increase in distance from nearest hazardous waste site increased property value by $1854.

Lewis Court is widely credited with formulating the idea behind the hedonic pricing method in 1941, but at least two applications of it were published in the 1920s (Colwell & Dilmore, 1999). Ronald Ridker first used housing data to estimate changes in environmental quality in 1967 (Freeman, 1993, pp. 367-370), and the method was refined by a cast of researchers from the mid-1960s to mid-1970s (Desvousges, Johnson, & Banzhaf, 1998, p. 17; Hanley & Spash, 1993, p. 74). The late 1970s through the 1980s then saw an "explosion" of empirical studies of the monetary values of nonmarket environmental amenities based on hedonic pricing (Freeman, 1993, pp. 367-370). The method has been used to value all sorts of natural resources including noise and vistas, but it has been most commonly used to value air quality (Desvousges, Johnson, & Banzhaf, 1998, p. 18).

Like the travel cost method, a major advantage of hedonic pricing is that it indirectly measures the values of nonmarket goods from real, observable market behavior (Desvousges, Johnson, & Banzhaf, 1998, p. 17). However, also like the travel cost method, one of its major disadvantages is its reliance on the assumption that consumers of the related market goods perceive the differences in environmental amenities, which might not be evident (David, 1971; Hite, 1998). For example, the quality of air between neighborhoods might not be readily apparent to home buyers. Also, consumers might not necessarily be able to select their most preferred bundle of amenities from the complete range of possible combinations because marketed goods like homes are "pre-packaged" (Freeman, 1993, pp. 415-416; Ofiara & Seneca, 2001, pp. 212-215). That is, each amenity of the marketed good might not be able to vary independently of the others, causing problems with multicollinearity. Another problem with the method, especially when applied to home buying, is that many amenities of the marketed good are only available in discrete units, not continuous variables (e.g., number of rooms, waterfront, etc.). This violates basic assumptions of the statistical model (Ofiara & Seneca, 2001, pp. 212-215), though it is still often applied in such cases. The method is also sensitive to the functional form selected for the hedonic price equation and to omissions of variables (Hanley & Spash, 1993, pp. 78-79). Finally, like other weak complementarity methods, hedonic pricing underestimates willingness to pay.

Contingent Valuation

The final method, contingent valuation, takes a very straightforward approach to valuing environmental goods. It directly asks people the maximum amount they would pay for an improvement in an environmental amenity, or the minimum amount of payment they would accept for a reduction in an environmental amenity, if it were traded in a market. The responses to the question then represent the value of the environmental change (Freeman, 1993, p. 165). In this method, values are not indirectly inferred from market behavior, but directly asked outside of a real market.

Data for contingent valuation are gathered through carefully structured surveys that describe hypothetical markets in which the goods are traded. The surveys generally include descriptions of the environmental good and the market in which it is hypothetically traded, questions about the values of the environmental good, and questions about the respondents' demographics and prior knowledge (Desvousges, Johnson, & Banzhaf, 1998, p. 14; Freeman, 1993, p. 170; Mitchell & Carson, 1989, pp. 3-4).

The valuation questions can take several different forms. One distinction is whether the questions ask for respondents' willingness to pay (WTP) for an environmental good, or their willingness to accept (WTA) payment to forgo the good[3]. Willingness to pay is the maximum sum of money a respondent would be willing to pay for an increase in an environmental amenity or to prevent a degradation in one. Willingness to accept is the minimum sum of money a respondent would require to forgo an improvement in an environmental amenity or to accept a degradation in one. WTP takes as its reference point the absence of the improvement while WTA takes as its reference point the improvement, and WTP is constrained by income while WTA is not (Freeman, 1993, p. 8).

Other distinctions that can be made about the types of valuation questions include the structures of the questions and the hypothetical market (Desvousges, Johnson, & Banzhaf, 1998, p. 15; Freeman, 1993, p. 165; Hanley & Spash, 1993, p. 55; Ofiara & Seneca, 2001, pp. 199-200). Valuation questions can be open-ended or offer discrete choices. Open-ended valuation questions allow respondents to state any value they choose. Thus the responses are continuous variables. On the other hand, discrete choice

[3] Willingness to pay for an environmental improvement and willingness to accept payment for an environmental degradation are both measures of compensating variation, one construct of consumer welfare. For a complete discussion of welfare measures in environmental valuation, see Freeman's (1993) chapter 3.

questions confine the respondents to choose one of multiple pre-determined values. Discrete choice questions can also take many forms including referendum, iterative bidding payment cards, etc. Responses to bidding games and open-ended WTP questions require no analysis to obtain a measure of individual welfare change (Freeman, 1993, p. 173), while responses to discrete choice questions do require aggregate analyses. The types of hypothetical markets that are commonly used include taxes, fees, and other mechanisms of public funding.

After responses to contingent valuation surveys are gathered, the data can then be analyzed to estimate mean WTP or WTA values (Hanley & Spash, 1993, p. 54), test hypotheses about the influence of income and other demographic characteristics on value (Freeman, 1993, p. 173), and estimate total benefits through extrapolation of results to the relevant populations (Mitchell & Carson, 1989, pp. 3-4).

An example of the contingent valuation method used to value an environmental good is Kramer and Mercer's study of the value of tropical rain forest protection (1997). They surveyed a random sample of 1200 United States residents in 1992 and asked them their willingness to donate to a United Nations fund to increase from 5% to 10% the amount of the world's rain forests preserved in their natural state. They concluded that Americans are willing to make a one-time donation of $21 to $31 per household. When extrapolated to the population of households this range becomes $1.9 to $2.8 billion.

The method was originally proposed by Robert Davis in 1963 (Mitchell & Carson, 1989, p. 9), and it underwent substantial refinement throughout the 1970s and 1980s (Hanley & Spash, 1993, p. 53; Mitchell & Carson, 1989, pp. 9-14). It gained approval by the federal government for valuing costs and benefits under the Water

Resources Council's 1979 "Principles and Standards for Water and Related Land Resources Planning" and under the Comprehensive Environmental Response, Compensation, and Liability Act of 1980 (Mitchell & Carson, 1989, p. 13). In 1993, it received a major endorsement when the National Oceanic and Atmospheric Administration (NOAA) proposed the first federal government guidelines for its use in environmental policy analysis (Rosenbaum, 1998, p. 168). The endorsement followed NOAA's use of contingent valuation in its assessment of natural resource damages caused by the *Exxon Valdez* oil spill in Prince William Sound. This high-profile use of contingent valuation, and the subsequent federal endorsement of it, helped make it a broadly accepted method of environmental valuation. It has been widely used to measure all kinds of environmental goods such as water quality, biodiversity, and wildlife.

Contingent valuation is unlike the hedonic pricing and travel cost methods in very important ways. First, it is able to measure nonuse values (Freeman & Kopp, 1999, p. 53; Ofiara & Seneca, 2001, pp. 200-204). These are values people place on environmental goods from which they gain no actual use (Freeman & Kopp, 1999, pp. 52-53; Mitchell & Carson, 1989, p. 63; Ofiara & Seneca, 2001, p. 36). Use values, on the other hand, are those derived from the current direct or indirect use of the amenity (Mitchell & Carson, 1989, p. 62). For example, in the case of the *Exxon Valdez* oil spill, Freeman and Kopp (1999, p. 51) explain that "People who have never been to Prince William Sound and never plan to visit nevertheless may feel a keen sense of loss over the damages it has sustained from ten million gallons of spilled oil." This sense of loss reflects the nonuse value placed on the condition of the place. Nonuse values are sometimes called "intrinsic values" or "existence values" to contrast it with instrumental value (Freeman, 1993, p. 142). But some insist that existence values are actually a type of nonuse value that stems from individuals' knowledge that the environmental amenity exists for future generations (Ofiara & Seneca, 2001, p. 37). The net nonuse value for an environmental amenity is its total value minus the use value (Mitchell & Carson, 1989, p. 69). In contrast to the contingent valuation method, the hedonic pricing and travel cost methods can only measure use values. Of course, some are skeptical of this advantage of contingent valuation specifically because no other method can measure it. These skeptics rightfully note that the values cannot be verified without actual payments made. (Ofiara & Seneca, 2001, pp. 200-204)

Second, data for contingent valuation come from *stated preferences* of respondents, rather than from market behaviors. The value a respondent places on an environmental good is assumed to be the statement made in response to the question. In contrast, the hedonic pricing and travel cost methods are based on *revealed preferences* from the subjects' real market consumption of weak complements of the environmental goods. Contingent valuation's reliance on stated preferences makes the method susceptible to strategic responses. Different circumstances and perceptions of respondents provide incentives to some to deliberately overstate or understate their true preferences (Carson, Flores, & Meade, 2001; Mitchell & Carson, 1989, pp. 143-146), or at least provide no incentive for respondents to respond truthfully (Freeman, 1993, pp. 167-168). For example, a contingent valuation study designed to value the protection of the Coastal Plain of the Arctic National Wildlife Refuge from oil and gas exploration might ask respondents their willingness to pay for its permanent protection. Environmentalists who do not expect to actually pay the amount they state could have an incentive to overstate their preference, especially if they believe the overvaluation would influence the actual protection of the area. On the other hand, oil and gas company employees or other local residents who stand to benefit from the potential increase in industrial activity could have an incentive to understate their preferences, with hope of influencing actual development of the area. In both cases, there is little incentive to answer truthfully, largely because it is a hypothetical market (Ofiara & Seneca, 2001, pp. 200-204). However, there is much evidence that suggests that these potential strategic responses are not actually a significant problem for contingent valuation studies in most circumstances (Mitchell & Carson, 1989, p. 170). Respondents apparently tend to answer truthfully.

Contingent valuation's reliance on stated preferences also makes it dependent upon the respondents' comprehension of the surveys and hypothetical markets. This could lead to problems including difficulties with answering the valuation questions, and the formation of preferences during the survey (Bjornstad, Cummings, & Osborne, 1997; Carson, Flores, & Meade, 2001; Desvousges, Johnson, & Banzhaf, 1998, pp. 15-16). That is, respondents might have trouble understanding the hypothetical market, and they might decide on their response during the survey instead of stating an established preference. Both of these problems stem from the fact that the described market is not real. Because nonmarket goods are not sold in the market, respondents have not had previous opportunity or need to evaluate their preferences for them.

Third, contingent valuation elicits "ex ante" measures of value, while the travel cost and hedonic pricing methods obtain "ex post" measures (Freeman, 1993, pp. 14-15; Ofiara & Seneca, 2001, p. 198). The ex ante measures are based on presumed future behavior reflecting the stated preferences, while ex post measures are based on market behaviors that have already occurred. This attribute of contingent valuation further makes it susceptible to strategic responses because stated behavior (i.e., willingness to pay) does not necessarily equal real future behavior. Indeed, Freeman (1993) prescribes more use of ex post analysis to validate ex ante analyses (p. 14).

Despite the vulnerabilities of contingent valuation, there is a growing body of research that suggests that properly designed and implemented contingent valuation studies can result in reliable and valid valuations of environmental amenities (Freeman & Kopp, 1999, p. 53).

Finally, economic theory predicts that welfare estimates from contingent valuation should be different than those from hedonic pricing and the travel cost method, but there is dispute over the direction of the difference. On one hand, as previously stated, the weak complementarity methods underestimate willingness to pay for environmental improvements. Because the enjoyment of a nonmarket environmental good is inseparable from the complementary market good, the willingness to pay for the environmental good cannot be directly established. Instead, maximum willingness to pay is estimated to be equal to at least the attributable portion of the amount paid for the market good. Contingent valuation avoids this bias by directly asking consumers for their *maximum* willingness to pay for environmental improvements. On the other hand, Brookshire, Thayer, Schulze, and d'Arge (1982) have argued that value estimates from hedonic pricing should be larger than those from contingent valuation[4]. They note that

[4] While their argument is specifically applied to the difference between hedonic pricing and contingent valuation estimates, the logic is applicable to the differences between travel cost and contingent valuation estimates as well.

hedonic pricing estimates are derived from the variation in preferences among many different consumers (i.e., multiple indifference curves between environmental amenities and income). In contrast, contingent valuation estimates are derived from constant preferences (i.e., indifference curves) of individual consumers. They demonstrate that the consequence is that the willingness to pay derived from static preferences (contingent valuation) must be less than or equal to that derived from varying preferences (hedonic pricing). Smith and Pattanayak (2002) make the same observation, identifying the difference as that between Marshallian demand measured by hedonic pricing and the travel cost method, and Hicksian demand measured by contingent valuation. Regardless of these differences, policy prescriptions for these three methods of valuation do not make such distinctions, as described in Chapter 2.

Beside the travel cost, hedonic pricing, and contingent valuation methods, other ways of valuing environmental goods have also been developed. But some are variations or combinations of these three, and none are as popularly used in environmental valuation. Contingent ranking, a derivative of contingent valuation, is a survey method in which respondents are asked to rank several goods in order of preference, to determine marginal rates of substitution. If one of the goods has monetary value, then the willingness to pay for an environmental good in the list can be calculated (Freeman, 1993, p. 166). Contingent activity is a combination of contingent valuation with the travel cost method, in which survey respondents are asked how they would change their level of some activity in response to a change in an environmental amenity (Freeman, 1993, p. 166). The averting-behavior method uses the household-production framework to estimate the value of environmental quality (Desvousges, Johnson, & Banzhaf, 1998, pp. 18-19). The restoration and replacement cost method values environmental changes through calculations of the costs needed to physically implement improvements. It is relatively easy to implement, but does not account for individuals' preferences for changes in environmental quality. It also does not guarantee that people are actually willing to pay the cost of improved environmental quality (Desvousges, Johnson, & Banzhaf, 1998, pp. 21-22). The random utility model (RUM) is a relatively recent, sophisticated extension of the travel cost method (Freeman, 1993, pp. 478-479). Instead of determining numbers of visits to a site as in the travel cost method, the RUM determines whether to visit a site and which site among multiple substitutes. The RUM thus estimates people's utility function by focusing on each separate choice of which sites to visit or whether or not to visit.

The development of these tools to value environmental goods in monetary terms has apparently filled the gap that had existed in cost-benefit analysis of policies affecting the environment. But these solutions have been the focus of much debate and skepticism, as discussed in the next chapter.

CHAPTER 2

LITERATURE REVIEW

Popularity and Controversy

All of these methods have facilitated the use of cost-benefit analysis in environmental decision making, and they have been used extensively in recent decades. But their use has been fraught with controversy. This chapter begins by describing the growing use of environmental valuation tools due to two main factors: the growth of cost-benefit analysis as a major decision-making tool in environmental policy, and the growth of benefit transfer studies as a short-cut to valuing environmental goods. It then surveys the controversies surrounding environmental valuation and describes the need for the assessment of the valuation tools.

History and Growth

The growing use of these environmental valuation tools parallels the popularity of cost-benefit analysis. The growth of cost-benefit analysis, in turn, has been due to several factors including the status of economics in the social sciences, the rational and empirical foundations of cost-benefit analysis and its common unit of measure, and extensive government use and prescription.

On a macro-level, the use of cost-benefit analysis and its tools have benefited from the dominance of economics in the social sciences. Rees (1994) explains that the economic paradigm dominant in the current worldview has its roots in 19th century scientific materialism, a "deep entrenchment of scientific rationality and its companion, social utilitarianism, as the primary beacons of human progress" (p. 438). He claims that economics strove to be to the social sciences what physics is to the natural sciences, the fundamental science to which all else reduces: "The founders of the neoclassical school, impressed with the spectacular successes of Newtonian physics, strove to create economics as a sister science" (p. 438).

Cost-benefit analysis is a primary evaluative tool in economics, and it exhibits characteristics of the scientific method's positivism that stemmed from the natural sciences. As Parsons (1995) explains, "The attractiveness of [cost benefit analysis] as a tool of decision-making is somewhat obvious: it provides an apparently neutral technique for identifying goals, their impacts, and their costs and benefits, and it creates a measurable, 'objective' statement which can serve to aid the formulation and selection of choices and options" (p. 410). This ability to sort out and rank competing alternatives makes the method popular with public decision makers, and provides them "the essential core of rational analysis in government decision-making and in the legitimation of decisions." Ellis (1998) adds that the method is also appealing because it is empirically verifiable. "It is an empirical fact that people value happiness, and there are even tangible ways to *measure* this value, even if only approximately and indirectly" (p. 57, original italics).

Cost-benefit analysis's unit of measure also contributes to its popularity (Patton & Sawicki, 1993, p. 210). The method typically accounts for the costs and benefits of alternative choices in monetary terms, which are well understood and can make competing alternatives commensurable (Freeman, 1993, p. 12). By converting values to dollars (or other applicable currency), analysts can compare alternatives and easily convey the results. Bardach (2000), in his policy analysis text, even prescribes this measure for nonmarket goods. "Using money as the metric [for nonmarket goods] is a very good idea, and it often works much better than one might imagine. For instance, even the 'value of life' can sometimes be reasonably well described by the metric 'willingness to pay X dollars for a reduction in the risk of death by Y percent a year' or something like it" (p. 38). Environmental valuation tools aim to measure these monetary values of nonmarket environmental goods.

Besides being well understood and commensurable, monetary measures also appear precise and quantitative, two additional qualities of choice for government decision makers. These qualities reduce the necessity, and the appearance, of bureaucrats making case-by-case judgments, and thus reduce the possibility of wide variances in decisions for similar situations (Ellis, 1998, p. 152).

Finally, the growth of cost-benefit analysis in public decision making, and the corresponding growth of environmental valuation tools, can also be attributed to their prescription and use by government. At the highest level of governance in the United States, most presidents since Gerald Ford have actively and formally promoted some form of cost-benefit analysis in government decision making (Ofiara & Seneca, 2001, pp. 58-59). Ford's Executive Order 11281, in 1974, required all federal agencies to quantify and publish costs and benefits of anticipated new standards in the form of inflationary impact statements. It was the first executive rule requiring the comparison of costs and benefits. In 1978, President Jimmy Carter issued Executive Order 12044, requiring economic impact assessments of proposed regulations. Agencies were required to select alternatives with the least economic burden. President Ronald Reagan followed with his

Executive Order 12291 in 1981 which returned to an emphasis on cost-benefit analysis and recommended actions whose benefits exceed costs. In 1993, President Bill Clinton issued Executive Order 12866 which relaxed Reagan's recommendation for net benefits in proposed actions, but still required cost-benefit analyses to be performed. Most recently, in 2002, President George W. Bush's administration has also prescribed and emphasized the use of cost-benefit analysis to evaluate the efficacy of environmental and other government regulations (InsideEPA, 2002).

At the bureaucratic level, promulgated rules and regulations have also prescribed cost-benefit analysis, specifically in environmental policy. In U.S. federal water resources planning, guidelines for economic assessment of water resources projects were established by the U.S. Water Resources Council in 1973 and 1983. Economic methods recommended for environmental valuation included the travel cost method and contingent valuation (Loomis, 1986; Ofiara & Seneca, 2001, pp. 56-58). The Department of the Interior and the National Oceanic and Atmospheric Administration (of the Department of Commerce) have also directly promoted environmental valuation techniques when they promulgated procedures for the assessment of natural resource damages. The Oil Pollution Act of 1990 and the Comprehensive Environmental Response, Compensation, and Liability Act of 1980 (also known as "Superfund") require the assessment of economic losses due to injuries to natural resources, and such losses can include nonuse values (Ofiara & Seneca, 2001, p. 36). The Department of Interior's rules recommend the use of several valuation methods for this task: market price, appraisal, factor income, travel cost, hedonic pricing, unit value, and contingent valuation methods (Ofiara & Seneca, 2001, pp. 70-71). The National Oceanic and Atmospheric

Administration's rules also recommend several techniques: habitat equivalency, travel cost, factor income, hedonic pricing, market models of demand and supply, contingent valuation, conjoint analysis, and benefits transfer methods (Ofiara & Seneca, 2001, p. 78). It is worth noting that the methods common to both agencies' rules include the travel cost, hedonic pricing, and contingent valuation methods. While these methods observe different things to infer the values of environmental goods, as discussed in Chapter 1, these regulations prescribe the alternative methods interchangeably, making an implicit assumption about their convergent validity.

Congressional legislation has also promoted cost-benefit analysis and environmental valuation. The Flood Control Act of 1936 is commonly seen as the start of systematic cost-benefit analysis in any area of U.S. federal policy (Hines, 1973). Much later, two environmental statutes required the *balance* of costs and benefits in setting environmental standards: the Federal Insecticide, Fungicide, and Rodentcide Act and the Toxic Substances Control Act (Van Houtven & Cropper, 1999, p. 43). Two other high-profile environmental laws, however, prohibit the use of cost-benefit considerations in establishing standards (Rosenbaum, 1998, pp. 158, 163). These prohibitions reflect a controversy in environmental valuation discussed in the next section, but also reflect the wide use of cost-benefit analysis in the government and the need to normatively assess its proper use. Indeed Van Houtven and Cropper (1999, 1996) examined whether these prohibitions or requirements for cost-benefit analysis make any difference in the regulations that were written, and found that both costs and benefits appear to influence the setting of standards, regardless of the mandates. More costly standards are less likely to be implemented, while standards that save more lives are more likely to be implemented. In all the environmental regulations they examined, cost-benefit considerations explained 85% of the standards set.

In 1994, the Republican revolution in the U.S. House of Representatives included an attempt to expand cost-benefit analysis to all social policy decisions in every federal agency (Ellis, 1998, p. 147), and to require proposed regulations to show projected net benefits before adoption (Rosenbaum, 1998, pp. 158-159). This would have been a significant expansion of Reagan's requirements. The proposal was defeated, but the debate illustrated the clout that cost-benefit analysis has gained in public policy decision making.

The growing use of environmental valuation techniques is not limited to the United States. European nations have traditionally placed much less emphasis than the U.S., both in law and in practice, on economic efficiency in environmental policy. However, the advent of the European Union will likely change this (Navrud & Pruckner, 1997). The Union has prescribed using environmentally adjusted national accounting systems (Repetto, 1998) and cost-benefit analysis in environmental policy. Both prescriptions would increase the use of environmental valuation studies there.

Finally, the growth of cost-benefit analysis and environmental valuation techniques is most obviously evident in the numbers of environmental valuation studies performed. Van den Bergh, Button, Nijkamp, and Pepping (1997) have called the growth a "renaissance of interest in the use of traditional cost-benefit techniques based on expressing environmental effects in common monetary units," resulting in "a considerable number of studies that have sought to place monetary values on negative environmental externalities" (p. 46). For example, Carson, Wright, Alberini, Carson, and

Flores (1994) compiled an extensive bibliography on just one method of environmental valuation, contingent valuation, and found 1,672 articles. Overall, these environmental valuation efforts have been used for cost-benefit analyses of government actions, environmental damage assessment, environmental costing, and environmental accounting (Navrud & Pruckner, 1997).

Benefit Transfer

The use of these tools to measure the values of environmental goods has also been popularized and highlighted by their use in benefit transfer studies. These are policy analyses that use results of existing valuation studies, implemented for their own specific purposes and contexts, to evaluate policy choices in other contexts (Brookshire & Neill, 1992). However, they are much more than a deductive application of a developed theory to a new circumstance. Rather, benefit transfer refers specifically to the transfer of values, especially environmental valuations from one situation to another (Desvousges, Johnson, & Banzhaf, 1998, pp. 3-5).

The basic framework of the benefit transfer method follows four broad steps (Ofiara & Seneca, 2001, p. 265; Kask & Shogren, 1994):

1. identify the good or nature of the commodity to be assessed,
2. identify potentially applicable studies,
3. evaluate their relevance to the transfer under consideration, and
4. develop benefit estimates based on the applicable studies.

The transfer of values from existing studies to the current cases involves adjustments based upon the differences between the contexts of the past and current studies. These differences could be demographic (e.g., differences in income, setting, etc.) or methodological (e.g., differences in valuation tools, environmental measures, etc.). Desvousges, Johnson, and Banzhaf (1998, pp. 5-9) prescribe specific procedures in making these value adjustments:

1. identify the cause and effect links between the variables of interest and the existing literature that quantify them;
2. obtain background information on the current case, such as baseline environmental quality and socioeconomic data;
3. combine the information from the above two steps to perform preliminary assessment of benefits or costs and identify new areas for adjustment, and make adjustments accordingly;
4. transfer the benefits and/or costs to the current case, quantifying the linkages; and
5. project the per unit results from step 4 to the relevant market to obtain the total benefits and costs.

Morgan and Owens (2001) provide a recent example of benefit transfer to value water quality improvements. Their study estimated the monetary value of benefits the Clean Water Act and other environmental policies have had on the Chesapeake Bay from 1972 to 1996. They do so by first comparing the actual conditions of the Bay in 1996 with estimates of its 1996 condition if these policies were not implemented. The difference was estimated to be a 60% improvement in water quality with the policies in place since 1972. They then selected existing valuation studies based upon the environmental good being valued (water quality), the site of valuation (Chesapeake Bay), and similar measures of water quality that they used in their modeling (nitrogen, phosphorus, and fish catch rate). They found two such studies. Both measured only recreational values, so the benefit transfer also only estimated recreational values. From these studies' results, the authors transferred values to estimate a 60% improvement. The result was an aggregate annual value of $357.9 million to $1.8 billion.

Application of the transfer method in the United States began concurrently with the early prescription of cost-benefit analysis with the Flood Control Act of 1936. Both cost-benefit analysis and the transfer method were also promoted for damage assessment by Superfund, which established liabilities for natural resource damages. The promulgation of its rules led to required compensation within the welfare economics paradigm (Desvousges, Johnson, & Banzhaf, 1998, p. 2).

Some of the appeal of the benefit transfer method are obvious. Most obvious are its economies of money and time (Desvousges, Johnson, & Banzhaf, 1998, pp. 9-10; Freeman, 1993, p. 484; Ofiara & Seneca, 2001, p. 264). Implementing original valuation studies can be quite expensive. When budgets do not allow for this, the benefit transfer method provides a means of applying results from past studies to current situations. Original studies are also labor intensive and can take much time to prepare and implement. When the luxury of time does not exist, as is often the case with public policy analyses in the political arena or damage assessments in judicial contexts, the benefit transfer method can produce estimates in significantly less time. Even when money and time are available for original studies, the need for a current valuation might not warrant a new application, and benefit transfer can produce acceptable estimates.

The method also has other, less obvious but nonetheless important, advantages: the method is consistent with economic theory (Desvousges, Johnson, & Banzhaf, 1998, p. 9); it organizes the underlying linkages between the variables, providing logical extrapolations of values; and its use can identify areas needing more valuation attention.

With these advantages, the benefit transfer method has become a “bedrock” (Desvousges, Johnson, & Banzhaf, 1998, p. 1) of environmental policy analysis, and the valuation studies supporting the method have become useful beyond their original intent. Indeed, the popularity of the benefit transfer method is evident in the development of benefit transfer databases. These databases catalog published environmental valuation studies, recording their conclusions as well as many demographic and methodological characteristics of the studies. Policy analysts conducting benefit transfer studies can then use these databases to search for original valuation studies pertinent to their new cases. Environment Canada, Canada's national agency concerning environmental policy, has sponsored a major endeavor to produce such a database called the Environmental Valuation Reference Inventory[1]. The U.S. Environmental Protection Agency has also funded efforts to create databases of environmental valuation studies.

Proponents and Opponents

Despite the popularity of these environmental valuation methods and their use in environmental cost-benefit analysis, they are controversial. Economists and others have taken all sides of the discussion on the reliability and validity of the methods and their normative implications. Table 2.1 summarizes the major arguments taken by proponents and opponents of environmental valuation techniques.

On one extreme, strong proponents of environmental valuation and cost-benefit analysis place much confidence in the theory behind the tools used. One such proponent is Freeman, who has published extensively on environmental valuation tools. He clearly prescribes the application of cost-benefit analysis to environmental issues and nonmarket goods as an ethical act (Freeman, 1997, p. 189) that provides "normative guidance"

---

[1] See http://www.evri.ec.gc.ca/evri/.

(Freeman & Portney, 1999, p. 16). The ethic is based on utilitarian efficiency (Ellis, 1998, pp. 15-32; Paris & Reynolds, 1983) applied to the management of scarce resources:

> "If society is to make the most of its scarce resources, it should compare what it receives from pollution control and environmental protection activities with what it gives up by taking resources from other uses… Society should undertake environmental protection and pollution control only if the results are worth more, in terms of individuals' values, than what is given up by diverting resources from other uses. This is the underlying principle of the economic approach to environmental policy. Benefit-cost analysis is a set of analytical tools designed to measure the net contribution of any public policy to the economic well-being of the members of society." (Freeman, 1997, p. 189).

Besides promoting the utilitarian ethic behind environmental economics, Freeman (1993) also promotes the valuation tools used in it. In his seminal text on environmental valuation, he describes the effectiveness of economic valuation tools for environmental goods by writing, "… I believe an economist could specify the economic theory and models he would use, the data he would like to have, and the empirical techniques he would apply to the data to obtain measures of benefits" (p. 484). Furthermore, he refuted visceral objections by critics who say nonmarket environmental goods such as ecology, health, and aesthetics cannot or should not be monetized.

> "It is not correct to say that there are some things like human health and safety or the preservation of endangered species that cannot be valued in terms of dollars or some other numeraire. The real world often creates situations where trade-offs between such things as deaths avoided and some other things of value cannot be avoided. The questions really are how the problem of making choices about such trade-offs is to be approached, and what information can be gathered to help in the problem of choice." (Freeman, 1993, p. 10)

Table 2.1: Controversy in Environmental Valuation

	Strong Proponents	**Marginal Proponents**	**Marginal Opponents**	**Strong Opponents**
Position	Valuation tools are valid and reliable.	Valuation tools have vulnerabilities, but monetization of environmental goods is necessary.	Monetary valuation is a good idea, but the state of the art is too problematic.	Utilitarian ethics are inappropriate for judging environmental policies.
Major arguments	Theory is valid, and methods are reliable, if not precise. Reflects real tradeoffs people make. Valuation is ethical and practical.	Some estimate of monetary value is better than none. Much caution must be used in implementing and interpreting results of valuation tools. There are no better alternatives.	Methods are neither objective nor reliable. The framework is blind to inequities.	Values people place on environmental goods are neither market-based nor utilitarian. Behavior does not follow economic theory.
Example opinion leaders	A.M. Freeman III, J. Loomis, NOAA panel.		M. Sagoff, H. Gintis, S. Kelman.	

Others have echoed this view. Costanza et al. (1997) argued that people express values for environmental goods every time they make choices and tradeoffs concerning them. Weaver (1996) provided some survey evidence that supports this claim. In his study of U.S. farmers, Weaver concluded that some people do think in utilitarian, and self-interested, terms when dealing with environmental issues. For Freeman and other leaders of environmental valuation, state of the art tools such as the travel cost method, hedonic pricing, and contingent valuation represent valid approaches to infer monetary value of environmental goods (Loomis & Walsh, 1986). Furthermore, utilizing these tools allows government decision makers to effectively address a wide range of environmental policy problems such as setting efficient levels of environmental standards and analyzing the net benefits of existing and proposed regulations (Freeman, 1997, pp. 189-191).

However, Freeman (1993) notes that there are limitations to the tools and framework. First, the utilitarian ethic of welfare economics focuses on the welfare of *humans*. He concedes that environmental cost-benefit analysis is inadequate to quantify ecocentric values (p. 485). Second, he also notes that economic efficiency is not the only, or even the primary, valid criterion against which to judge environmental policies (pp. 8-9). As such, cost-benefit analysis should not be used as a simple decision rule replacing the judgment of decision makers. Rather, its conclusions should be weighed along side other criteria such as equity and acceptability. Third, despite his confidence in the valuation tools, the state of the art is not perfect. He states that the economic values people place on nonmarket goods can "seldom be measured with precision" (pp. 191-192), and that the political contexts in which cost-benefit analyses are performed can influence the outcomes of the analyses. That is, the tools are vulnerable to manipulation.

Richard Carson similarly defends the specific method of contingent valuation against criticisms of the method's reliability and validity. Carson, a leading promoter of the method (Mitchell & Carson, 1989), recently reviewed the major criticisms against contingent valuation and the evidence of the criticisms. He concludes that the problems seen in the results of contingent valuation studies are not due to the method itself, but instead are due to spurious factors in the design and implementation of specific studies. The alleged problems – such as strategic bias, scope effects, hypothetical bias, etc. – can be resolved with careful and proper design and implementation (Carson, Flores, & Meade, 2001). Furthermore, when accounting for these spurious factors, he finds that the results are consistent with economic theory. There is certainly empirical evidence supporting these claims. Many studies have been conducted that successfully alleviated problems including hypothetical and payment vehicle biases (e.g., Champ & Bishop, 2001; Cummings & Taylor, 1998; Morrison, Blamey, & Bennett, 2000), reliability (e.g., Carson et al., 1997; Reaves, Kramer, & Holmes, 1999; Sanders, Walsh, & McKean, 1991; Whitehead & Hoban, 1999;), and protest responses and yea-saying (e.g., Blamey, Bennett, & Morrison, 1999).

Even with the methods' limitations, strong proponents note that measuring the values of environmental goods has helped shape real policy decisions regarding the environment. For example, returning to the case of the *Exxon Valdez* oil spill in Prince William Sound, the accident caused the deaths of over 1600 birds and 700 otters, plus unknown thousands of other animals and plants (Keeble, 1999, p. 185). Government valuation of the environmental damages, plus estimates of the losses to the local economy, totaled $2.8 billion (Freeman & Kopp, 1999, pp. 51-53), which the government sought from Exxon. In October 1991, Exxon and the U.S. and Alaskan governments agreed on a settlement of $1.025 billion (U.S. Senate Committee on Energy and Natural Resources, 1999, p. 3). This money was designated for remedial and compensatory clean up, and research on spill prevention and remediation, under the approval of a federal and state trusteeship. Ultimately, the monetization of environmental goods has helped facilitate the restoration of the damaged ecosystem.

Less enthusiastic than Freeman but still supportive of environmental valuation are proponents who are concerned with the vulnerabilities of the tools but concede the necessity of measuring the monetary values of environmental goods in a market economy. The National Oceanic and Atmospheric Administration's blue ribbon panel on contingent valuation exemplify this position. The panel, which included Nobel Prize laureates Kenneth Arrow and Robert Solow, endorsed contingent valuation as a means of valuing environmental goods, but they tempered their endorsement with a long list of procedural cautions stemming from the technical vulnerabilities of the method (Arrow et al., 1993). They "believe" in the method, but conceded that it is a "brand new science [that] is fraught with pitfalls" (U.S. House of Representatives Committee on Merchant Marine Fisheries, 1991). Goodland and Ledec, from the World Bank's Department of Environmental and Scientific Affairs, are much more pragmatic in their endorsement of environmental cost-benefit analysis:

> "Despite the many deficiencies of CBA, it can still be useful for advancing environmental goals. Even unreasonably low or highly inaccurate estimates of environmental benefits and costs are better than none, because the alternative is to assume implicitly that these benefits and costs are zero. Rather than abandon CBA, environmentalists should insist that it take environmental and other social costs explicitly into account." (Goodland & Ledec, 1994, p. 451).

Other economists and policy analysts agree (Bardach, 2000, p. 39; Leonard & Zeckhauser, 1998, pp. 6-9). Rosenbaum's (1998, p. 165) list of cautions that should be taken when implementing cost-benefit analysis in environmental policy exemplify the concern of these marginal proponents. Among his prescriptions are the following:

- cost-benefit analysis does not need to be done for all environmental regulations, but it also should not be categorically excluded, because there are instances in which it can identify more efficient solutions;
- cost-benefit analyses should be open for review and challenge, since the method is vulnerable to manipulation; and
- Congress should specifically state in regulatory legislation the weight given to economic criteria such as costs and benefits, compared to other criteria of evaluation.

On the other side of the debate are a range of opponents to cost-benefit analysis applied to environmental policies and environmental valuation itself. Marginal opponents are those who like the idea of including monetary valuations of the environment in policy analyses, but see too many problems with the methods currently available (e.g., Green & Tunstall, 1991; Kellert, 1984; Neill, Cummings, Ganderton, Harrison, & McGuckin, 1994). Much of the criticism of environmental cost-benefit analysis focuses on a false sense of objectivity it projects. Such criticism notes that the rational and quantitative framework of cost-benefit analysis can lead analysts and the public to believe that its results are objective, value-neutral facts. However, cost-benefit analyses are often used to justify decisions already made or enhance the appearance of a favored alternative, rather than objectively compare alternatives (Paris & Reynolds, 1983; Tong, 1986, pp. 14-15). Findley and Farber (1992) have found evidence for this claim in federal Environmental Impact Statements. "Agencies often engage in extremely slipshod cost-benefit analysis of their proposals, frequently biasing their results in favor of their projected course of action" (p 52). Parsons (1995) summarized this criticism in a sentence: "Value-free it might look, but value-free it ain't" (p. 401).

Another criticism of cost-benefit analysis applied to environmental policy is its "ethical poverty" (Gillroy, 1992). The ethical foundation of cost-benefit analysis lies in efficiency, but there are other moral principles – such as equity, freedom, benevolence, etc. – that may be more applicable in public issues such as environmental policy (Sagoff, 1997). Kneese (1999, p. 55), a self-described long-time student and practitioner of cost-benefit analysis, agrees. He states that cost-benefit analysis cannot solve certain questions and decisions of environmental policy and may actually obscure them.

Of the alternative ethical principles, equity is an often-cited rival to efficiency in environmental issues. Cost-benefit analysis is blind to inequity (Ellis, 1998, p. 62; Paris & Reynolds, 1983) or, as proponents of it describe it, is "distributionally neutral" (Freeman & Portney, 1999, p. 17). Who bears the costs and who gains the benefits of any environmental decision is not of concern in a cost-benefit analysis, only the net balance to society is measured. Thus, critics note, even an efficient policy in terms of net benefits can lead to unjust outcomes. It may be economically efficient, for example, to build hazardous waste facilities in economically depressed areas, but it can also be unjust to the residents in the area. In this example, the benefits of properly handling hazardous waste is spread to the entire community while the burden of living near the facility is borne by a few.

Marginal opponents also note deficiencies in the environmental valuation tools that proponents of the tools also recognize. Some highlight the cognitive difficulties of answering valuation questions in contingent valuation (Desvousges, Johnson, & Banzhaf, 1998, pp. 15-16). Others point to the inability of revealed preference methods to include nonuse values in their measures, as well as the unverifiable and unsatisfactory state of the art of the one method that can measure it (Goodland & Ledec, 1994, p. 452). Some argue that even contingent valuation does not reliably measure nonuse value (Cummings & Harrison, 1995; Shechter, Reiser, & Zaitsev, 1998). Rosenbaum (1998), who supports cautious use of cost-benefit analysis in environmental policy, concedes that the tools' methods for valuing environmental goods follows "a tortuous, inevitable contentious logic" (p. 166).

Perhaps the most common criticisms of current environmental valuation tools are the inconsistencies and biases in their results (e.g., Balistreri, McClelland, Poe, & Schulze, 2001; Boyle, MacDonald, Cheng, & McCollum, 1998; Kealy, Dovidio, & Rockel, 1988; Loomis, Brown, Lucero, & Peterson, 1996; Loomis, Brown, Lucero, & Peterson, 1997; Seip & Strand, 1992). For example, the U.S. Environmental Protection Agency conducted a valuation study to estimate the value of benefits from improved visibility at the Grand Canyon due to scrubbers installed at electric utility facilities. Their estimate of annual benefits was $130 to $250 million. The electric industry conducted their own valuation study on the same issue and concluded an annual benefit of $50 million (Reisch, 2001, p. 5). The *Exxon Valdez* case is another example. NOAA conceded that the valuation studies of the damages to Prince William Sound varied by $5 billion (House of Representatives Committee on Merchant Marine and Fisheries, 1991). What are the consequences of such inconsistencies? King (1998) argues that the results of environmental valuation efforts can be so unreliable and subjective that their use in policy decision making has been detrimental to the environment instead of helpful. He reviewed cases of wetlands valuations and the policy decisions resulting from them and found that valuation efforts have led to the overuse and degradation of wetlands.

Finally, at the other extreme of the continuum, there are strong opponents of environmental valuation who object to the very idea of monetizing environmental values. Their arguments are mostly normative, pointing to the inappropriateness of utilitarian ethics applied to environmental policies. Kelman (1998), for example, argues that environmental goods are priceless and not for sale, so cost-benefit analysis is not applicable. Instead, he claims, environmental policies are more appropriately judged by deontological criteria.

Sagoff (1997, 1981) is a little more tempered in his criticisms, but still sides with more deontological approaches[2] to evaluating environmental issues. He grants that environmental valuation findings represent *some kind* of information, but not necessarily the true value of environmental goods. The true value people place on environmental goods, he argues, is not confined to market prices nor to people's desires as consumers. Instead, when people judge the value of environmental goods they do so from their obligations as citizens. "We act as consumers to get what we want fo*r ourselves*," he argues, but "we act as citizens to achieve what we think is right or best *for the community*" (Sagoff, 1981, p. 445, original italics). These two roles are not the same and Sagoff provides several examples of how people's preferences as consumers are not consistent with their judgments as citizens. More recently, Sagoff (1998) has

[2] Deontology is a theory of ethics based upon duties and rights. It stands in contrast to economics' utilitarian ethics, which is based on the greatest amount of benefits for the greatest amount of people.

acknowledged some utility in valuation exercises. Specifically, he developed a modification of contingent valuation that uses a jury-like group of respondents to collectively deliberate, and come to consensus, on some environmental value. But he maintains that the results of such deliberation reflect public and communal interests of citizens, not private preferences of consumers for which proponents of environmental valuation strive. To Sagoff, measuring private preferences in environmental valuation remains invalid.

There is some empirical evidence supporting Sagoff's claims. Gintis (2000) reviewed empirical studies evaluating choice behaviors and motivations and found that choices people make in many contexts often do not resemble the rational actor of neoclassical economic theory, whom he calls *homo economicus*. Instead, people are much more cooperative and less self-interested than economic theory assumes. The implication, Gintis suggests, is that an economic framework for environmental decision making is inappropriate. In an early cross-cultural study of the contexts of cost-benefit analysis, Smith and Hogg (1971) found that while the economic, consumer model of behavior may be culturally acceptable in the U.S., other cultures value being benefactors more than being beneficiaries. Cost-benefit analysis does not account for the possibility of willingness to incur great costs.

In psychological evaluations of contingent valuation, researchers have concluded that responses are often dominated by value expressions much broader than the intended appraisals of economic tradeoffs (Blamey, 1998) and that respondents object to the idea of specifying a monetary, utilitarian value for environmental goods (Hanley, Spash, & Walker, 1995; Jakobsson & Dragun, 2001; Jorgensen, Syme, Bishop, & Nancarrow, 1999). One study specifically tested Sagoff's hypothesis of the citizen-consumer dichotomy and not only found evidence that the dichotomy exists, but also concluded that people's preferences for environmental goods are not consistent with transitivity, an economic axiom on which environmental valuation rests (Common, Reid, & Blamey, 1997). The latter conclusion has been found in other studies also (e.g., Giraud, Loomis, & Johnson, 1999; Johnston & Swallow, 1999). One study measuring the value of wildlife with contingent valuation summarizes these concerns well:

> "…many respondents were either uncertain about their valuation, believed that wildlife should not be valued in dollar terms, or felt that the money should come from somewhere else (taxes and fees). Moreover, a majority of those who would pay exhibited behavior which appears inconsistent with the neoclassical economic theory of trade-offs between money and wildlife. In essence, this [contingent valuation study] may have asked people to choose between ordinary goods (income) and a moral principle… Such choices are likely to produce conflict and ambivalence, and the resulting behavior (protest and avoidance) is likely to be inconsistent with the usual economic assumptions… We therefore conclude that benefit-cost analysis should not be used to make decisions about wildlife recovery programs" (Stevens, Glass, More, & Echeverria, 1991, p. 333).

In a broader, sociological study, Kempton, Boster, and Hartley (1995) conducted an extensive survey to measure environmental values in America. They note that historians such as White (1967) describe Americans as viewing nature in utilitarian terms, due to a Judeo-Christian interpretation of Genesis as a license to use and master nature as they see fit. Such a view would reveal itself in utilitarian arguments for environmental protection or exploitation. Contrary to this, however, the authors found in their research that only a small minority of Americans take a purely utilitarian view. This was true of all five clusters in their sampling plan, representing a range of environmental perspectives from Earth First! members to sawmill workers. Instead, all clusters favored more aesthetic and biocentric perspectives (p. 102-115).

Besides these ethical arguments, some strong opponents also summarily dismiss the validity of environmental valuation efforts. Most economists deny the possibility of making cardinal or interpersonal comparisons of utility (Paris & Reynolds, 1983), yet environmental valuation tools attempt to do exactly these two things: measure individual utility and compile them. Others say the public nature of environmental issues make individual preference measurement inappropriate (Moody, 1974; Sagoff, 1994).

As if the above controversy over the validity and ethics of environmental valuation were not enough, the debate’s fire has been recently fanned by the growth of benefit transfer studies. As previously described, analysts using this method calculate environmental values not through original application of one of the valuation methods, but through review of existing literature that they deem most similar to the new circumstances they face. This approach assumes the generalizability of the existing literature (Smith, 1992), but does not necessarily verify it. This is problematic because most applications of the travel cost, hedonic pricing, and contingent valuation methods have been used to evaluate specific, local environmental changes, with their generalizability only qualitatively assessed or left to the judgment of their readers altogether. Indeed, there is generally little said about the generalizability of each application's results.

A few studies have attempted to test the reliability of benefit transfers by comparing their results with original applications of valuation methods. The results of these case studies have been mixed. Piper and Martin (2001) and Chestnut, Ostro, and

Vichit-Vadakan (1997) concluded that benefit transfer can produce reasonable estimates when using broad based environmental valuation models or when measuring values as portions of income rather than absolute amounts. Others have concluded that benefit transfer results are unreliable (e.g., Boyle, Bergstrom, & Poe, 2001; Brouwer, 2000; Downing & Ozuna, 1996; Kirchoff, Colby, & LaFrance, 1997; Loomis, Roach, Ward, & Ready, 1995). Brouwer's (2000) test resulted in transfer errors up to 475% of the estimates from original valuation applications.

Need for Evaluation

The wide use of the three environmental valuation methods, along with their controversy, necessitate their evaluation. On one hand, from a purely pragmatic perspective, environmental valuation methods are worthy of assessment simply because they are popular and exert real influence on policy (Rosenberg, 1992; Smith, 1993). Thousands of articles on these methods have been published in the last few decades, affecting the public policies and legal cases they analyze. In the *Exxon Valdez* case alone, environmental valuation influenced a billion dollar settlement.

On the other hand, from both academic and pragmatic perspectives, environmental valuation methods are in need of assessment because there is much controversy over their reliability and validity. Besides the methodological and normative concerns of the proponents and opponents discussed above, other academic concerns include the high variance of values for similar goods valued by different methods or under different circumstances. This wide variance from different studies is "regularly cited" but little effort has been made to evaluate the causes (van den Bergh, Button, Nijkamp, & Pepping, 1997, p. 46). The effects of valuation studies' characteristics – such as ages, valuation methods, and environmental goods – on their results must be determined in order to appropriately use them in future benefit transfer studies (Ofiara & Seneca, 2001, p. 265; Smith, 1992). Also, the increase in the use of benefit transfer studies has demanded development of standards and protocols for their implementation (Boyle, Bergstrom, & Poe, 2001; Desvousges, Johnson, & Banzhaf, 1998, p. 3). Finally, there is also plain public skepticism about the methods and their results. In the legal war following the *Exxon Valdez* oil spill, the controversy over the values of the lost wildlife and damaged ecosystem was due not only to the accuracy of the biological surveys, but also to the methods used to measure the values of the damages (Keeble, 1999, pp. 197-199). The public has become increasingly distrustful of "expert" opinions such as those given in cost-benefit analyses (Freeman & Portney, 1999, p. 19).

In short, the practical use of environmental valuation methods make their examination relevant. A well designed and implemented examination, in turn, could address some of the concerns described. For example, evaluative efforts could determine the convergent validity of the methods, and identify variables explaining the variances. Such findings could help refine the proper uses of the methods in original studies as well as in benefit transfer studies (van den Bergh, Button, Nijkamp, & Pepping, 1997, p. 52).

Given the popularity of these methods, the controversy surrounding them, and the potential refinements to be gained from their evaluation, it is surprising that there have been very few meta-analyses of valuation studies (van den Bergh, Button, Nijkamp, & Pepping, 1997, p. 46). A few have been performed, but none with the scope and approach done here. Existing meta-analyses have tended to focus on specific goods such as outdoor recreation (Rosenberger & Loomis, 2000; Shrestha & Loomis, 2001; Smith & Kaoru, 1990), wetland services (Woodward & Wui, 2001), ground water values (Boyle, Poe, & Bergstrom, 1994), and air quality (Smith & Huang, 1993). Some meta-analyses have focused on specific valuation tools (Smith, Bruford, & Wayne, 1996; Smith & Huang, 1993; Smith & Kaoru, 1990). However, these studies attempted to find the *determinants* of value, rather than test the validity of the methods. Indeed, Smith and Pattanayak (2002) reviewed 15 meta-analyses of nonmarket valuation and found that 13 of them, including some of those cited here, sought to synthesize valuations of specific goods and identify the determinants of the values.

One meta-analysis, however, attempted to address the validity of two valuation methods. Carson led a meta-analysis of studies that tested the convergent validity of values measured by both contingent valuation and revealed preference methods (Carson, Flores, Martin, & Wright, 1996). His findings generally support the validity of the methods, but they were based only on studies that compared stated and revealed preference methods together (i.e., the purpose of those original studies was to test the validity of contingent valuation). No one has yet compared the results of *different* stated preference and weak complementarity studies for the purpose of testing the convergent validity of the valuation tools.

Also, the existing meta-analyses have used the outcome measures of the original studies (e.g., mean dollar values) when comparing results across the studies. This approach limits the scope of a meta-analysis to just those studies using similar measures, and it does not account for the variability in the results within each study. Both of these factors limit the scope and significance of the conclusions made by such a meta-analysis.

Meta-analyses in other fields, however, typically account for the variability in the results of studies by calculating effect sizes as the outcome measure. As discussed in Chapter 3, effect sizes are standardized measures of study outcomes that not only make it possible to compare results from studies using different outcome measures, but also make such comparisons more meaningful by accounting for the variability in the original studies. These two characteristics of effect sizes allow the researcher to include more studies in a meta-analysis and to measure broader trends among the studies.

This dissertation research fills these gaps in the meta-analyses of environmental valuation by testing the convergent validity of three valuation methods – using the effect size approach – and refining their proper contextual uses.

CHAPTER 3

METHOD

The guiding questions of this research are, “are the results from different environmental valuation methods collectively reliable, are there contexts in which they are more reliable than others, and – to the extent possible – are the results valid?” My goal then is to refine the proper contextual uses of these three methods in original applications and benefit transfers. To meet this goal, there are two objectives for this research. First, it tests the convergent validity of the valuation methods. Convergent validity in this context is the measure of how closely results from different methods come together to one value (Bishop, Champ, Brown, & McCollum, 1997; Carson, Flores, & Meade, 2001; Whitehead, 1995). While convergence of results is not a sufficient determinant of validity, it is a necessary condition. Thus, this research can signal problematic results, but not affirm valid ones. While measuring convergent validity, this research also identifies moderating variables explaining the variance in the valuations. Second, this research folds the results of these statistical analyses into the broader normative discussion surrounding environmental valuation.

This chapter begins by introducing the meta-analytic framework that was used to assess the valuation tools and by describing the significance of the analysis in environmental policy. It then describes in detail the hypotheses of the research, the procedures and scope of data collection, and the statistical analyses performed with the data. Next, it addresses some difficulties in the research that were anticipated and how they were addressed.

Meta-Analysis of Environmental Valuation

To address the controversies surrounding these environmental valuation methods, I performed an extensive meta-analysis of environmental valuation studies. Meta-analysis is a statistical method that integrates research findings from many studies addressing the same subject (Rosenthal & Rosnow, 1991, p. 491). By doing so it can assess the commonalities and variations across a range of prior studies. The method has significant advantages over traditional methods of research synthesis (van den Bergh, Button, Nijkamp, & Pepping, 1997, pp. 37-38). Unlike early, qualitative comparisons of valuation studies (e.g., Bergstrom, 1990; Gregory, Mendelsohn, & Moore, 1989; Loomis, 1987; Schulze, d'Arge, & Brookshire, 1981), meta-analysis synthesizes results quantitatively and thereby adds more information to descriptive taxonomies. The method is also statistically sound, unlike earlier quantitative methods such as vote-counting[1]. The result is a method that can provide a greater amount and quality of information than other methods of research synthesis. The statistical mechanics of meta-analysis varies with the information available from the studies evaluated, but the framework of all meta-analyses is common (Cooper & Hedges, 1994, pp. 9-13; Hunt, 1997; Hunter & Schmidt, 1990). It consists of collecting studies applicable to the research question, coding information from them, and analyzing the coded data.

The idea of quantitatively combining the results of different studies has been traced back to the early 20th century (Cooper & Hedges, 1994, p. 5; Rosenthal & Rosnow, 1991, p. 491), but the formal introduction of meta-analysis is widely credited to

[1] The vote-counting method of research synthesis resolves conflicting results from different studies by tallying the number of studies finding each result. The result with the most votes is favored. This approach ignores statistical explanations for differences in results, such as sample sizes and representativeness.

Gene Glass in 1976 (Cooper & Hedges, 1994, p. 5; van den Bergh, Button, Nijkamp, & Pepping, 1997, p. 35). Prior to this, quantitative syntheses of studies were rare (Cooper & Hedges, 1994, p. 6). Following Glass's introduction, which was in education research, the method gained popularity in psychology and medicine. However, its application in economics and environmental valuation has been "extremely limited" (van den Bergh, Button, Nijkamp, & Pepping, 1997, p. 38) even though some have noted its great potential to summarize information in these fields (Desvousges, Johnson, & Banzhaf, 1998, pp. 28-30; Skalski, 1995). Indeed, meta-analysis is applicable to any field with significant numbers of empirical studies with wide ranging or conflicting results. Environmental valuation is such a field.

Anticipated Results

Meta-analysis is used here to produce four important insights. First, at the most basic level, it provides descriptive statistics of existing environmental valuation studies. This includes the distribution of studies by environmental good, geography, time, and publisher. Such information helps identify gaps in the existing body of valuation studies and direct future studies (van den Bergh, Button, Nijkamp, & Pepping, 1997, pp. 40-42; Lipsey, 1994, pp. 116-117).

Second, a meta-analysis measures the central tendencies and variances of the values of environmental goods. For results that show wide variation, the meta-analysis tests for potential moderating variables. This is a key benefit of meta-analysis and is the most significant result of this research (Lipsey, 1994, pp. 118-120; van den Bergh, Button, Nijkamp, & Pepping, 1997, pp. 40-42). Moderating variables – sometimes called confounding variables outside of meta-analyses – are those that account for a significant portion of the variation in the primary results. By identifying moderators, this meta-analysis refines the uses of these environmental valuation tools. One important potential moderator that this study tests is the method of valuation. This constitutes a test of convergent validity. If the valuation method is a moderator then it can also show how the method affects the variation. For example, it could show which methods produce higher or lower estimates, which produce wider ranges, etc. Other important potential moderators tested include the type of environmental good (e.g., air, water, wildlife, etc.), and the measures of these goods. One would reasonably expect both of these factors to be moderators of value, but Chapter 2 describes the controversy over the ability to measure them. Confirming the effect of these potential moderators lends statistical evidence to the methods' validities. A list of potential moderators tested are discussed later in this chapter.

Next, the above two results lead to the third: this meta-analysis refines the defensible contextual uses of the three environmental valuation tools. On one extreme, if the analysis shows high convergent validity, then the refinements would not be any more confining than the prescriptions in the existing literature. On the other extreme, if the analysis shows low convergent validity, then the prescriptions would be more confining. In the middle are moderate results that could show some convergent validity, but also identifies moderators. In this case, this meta-analysis would detail the conditions in which different valuation tools produce the most defensible results.

Finally, with any of these above possibilities, this meta-analysis weighs heavily on the broader normative debate over environmental valuation, by providing statistical evidence to the basic question, "*Can* consumers reliably monetize the values they place on environmental goods?" If they can, we are still left with the broader, normative question, "*Should* we place monetary values on environmental goods?" But if they cannot, then the normative question becomes moot[2].

Significance

Two important questions that any meta-analysis is poised to address are "How confident can we be that the findings can be generalized beyond a small subset of populations, settings, and procedures?" and "Does the research advance the theoretical understanding of the phenomenon?" (Hall, Tickle-Degnen, Rosenthal, & Mosteller, 1994, pp. 18-21). Indeed, this meta-analysis has important practical and theoretical implications for environmental valuation.

Its impact on the practice of environmental valuation starts with estimating the economic values of environmental changes from the pool of existing studies for each good. It provides more broad estimates of the overall central tendencies and dispersions among many studies. More importantly, however, this meta-analysis tests how dependent the findings are on specific circumstances of the studies. One reason for the variation in results of environmental valuations is the "complex multidimensional nature of societal and environmental systems and interactions" (van den Bergh, Button, Nijkamp, & Pepping, 1997, p. 10). This meta-analysis helps sort out that complexity.

---

[2] Strong opponents of environmental valuation would object to this sequence of questions, claiming that the first question to ask is the normative one. They might argue that if monetization of environmental values should not be done, then the practical question is moot. However, I have chosen to ask the practical question first, to test the convergent validity of the methods from their own theoretical framework of econometrics.

The analysis clarifies the generalizability and utility of the valuation methods by identifying moderating variables that account for any trends in the variance of values. Comparison of these results among the three valuation methods also determines their relative strengths and weaknesses. The practical implications of such findings are obvious: they refine the defensible uses of environmental valuation in cost-benefit analysis and benefit transfer studies (Boyle & Bergstrom, 1992; Brouwer, 2000; Desvousges, Naughton, & Parsons, 1992; Walsh, Johnson, & McKean, 1992). Such findings are important to the many agencies and organizations that currently depend on such studies in their environmental decision making. They include the principle trustees of federal lands in the United States (the Department of the Interior [including the Fish and Wildlife Service, the National Park Service, and the Bureau of Land Management], the Forest Service, the National Oceanic and Atmospheric Administration, the Department of Defense, and the Department of Energy), the chief implementers of environmental legislation (such as the Environmental Protection Agency, the Fish and Wildlife Service, and the Army Corps of Engineers), as well as private organizations concerned with the same issues (for example, environmental groups and industry representatives).

This research also contributes to the understanding of the theoretical relationships between changes in environmental goods and their monetary values. Foremost, it determines whether the relationship can be measured reliably by different methods. If not, then opponents of environmental valuation will have a statistical argument to add to their thus far normative and qualitative arguments. If so, then proponents will have the statistical evidence to support their theoretical arguments.

Additionally, the significance of this research also stems from its scope. Few researchers have performed secondary analyses of environmental valuation, and none have done so with the scope done here. Unlike existing meta-analyses, this research compares the results of different stated preference and weak complementarity studies, covers a broad range of environmental goods, and collects a large sample of studies. The large scope gives the results of this research practical and theoretical impacts regardless of the actual results.

Hypotheses

After measuring the central tendencies and dispersions of the values of environmental goods, I then tested select variables for their moderating effects on the variance in the values. Stated formally in the format of statistical hypothesis testing, the null and alternative hypotheses are as follows:

H_0: there are no moderators.
H_1: there are moderators.

Moderators

First, the type of environmental good was tested as a moderating variable. Common sense and economic theory both state that different environmental goods have different values, and some tests have presented empirical evidence to this effect (e.g., Smith, 1996). To confirm this, alternative hypotheses were tested:

Potential moderator 1: type of good
H_0: type of good is not a moderator.
H_1: type of good is a moderator.

Beyond confirmation of what is intuitively obvious, this test could detail the degree of differences among environmental goods and estimate the variance of each. For example, perhaps environmental goods that do appear in some markets and are more familiar to consumers (e.g., land, water) exhibit less variance than those that are not found in markets (e.g., air, wildlife). Also, analysis of this hypothesis while controlling for the method of valuation revealed whether certain goods are better measured by certain methods (Freeman, 1993, p. 487).

Second, the magnitude of the environmental change was tested as a moderating variable. Again, common sense and economic theory appear to settle this question easily: the greater the change in environmental good, the greater the change in value (Freeman, 1993, p. 168; Freeman, 1998, pp. 26-28).

Potential moderator 2: magnitude of change
H_0: magnitude of change is not a moderator.
H_1: magnitude of change is a moderator.

But testing this hypothesis was not easy because environmental valuation studies have inconsistent measures of environmental change. Hedonic pricing studies, for example, tend to measure changes in environmental goods incrementally, such as the improvement in air quality in incremental units of pollution (e.g., parts per billion of particulate matter). Meanwhile, contingent valuation studies tend to measure changes in discrete, holistic terms, such as the improvement of air quality to a threshold condition (e.g., from "non-compliant" to "compliant" with air quality standards). Thus, this meta-analysis only gauges the effect of broad measures of magnitude, which are described in Chapter 4.

Third, and closely related to the magnitude of changes, is the method of measuring and describing the changes. Specifically, quantitative versus qualitative descriptions of the changes in the environmental goods were tested for their effect on values.

Potential moderator 3: type of description of environmental change
H_0: type of description is not a moderator.
H_1: type of description is a moderator.

Intuition does not provide clear guidance here. On one hand, quantitative descriptions would seem to produce more accurate estimates of value than qualitative descriptions. For example, the price one would be willing to pay for a television with a 19" screen would probably be easier to estimate than the price one would pay for a television with a "small" screen. However, quantitative descriptions of environmental goods are often obscure to a layperson, while qualitative descriptions are better understood. For example, people might better understand the difference between "bad" and "good" water quality than the difference between water with 100 or 30 "parts per million total suspended solids" even though the latter is one way water quality is quantified. Thus, it is not unreasonable to suspect that qualitative descriptions might produce more reliable estimates.

Finally, convergent validity of the three methods was tested (Freeman, 1993, p. 176; Lipsey, 1994, p. 119; Ofiara & Seneca, 2001, p. 265; van den Bergh, Button, Nijkamp, & Pepping, 1997, p. 50). That is, the method of valuation was tested as a moderating variable. If all three methods validly measure the same environmental values, then they should not moderate value. At least one limited meta-analysis has drawn this conclusion from an analysis of studies using both contingent valuation and revealed preference methods (van den Bergh, Button, Nijkamp, & Pepping, 1997, p. 50). Also, as discussed in Chapter 2, public policies prescribe these valuation methods without distinction, making an implicit assumption that values derived from these methods are comparable.

Potential moderator 4: method of valuation
H_0: valuation method is not a moderator.
H_1: valuation method is a moderator.

Yet, there are good reasons for skepticism. First, each of the methods has its own limitations which may bias their measures. Contingent valuation, for example, is especially prone to biases because it relies on the ability and willingness of respondents to answer truthfully. Also, the travel cost method is generally unable to account for substitution among sites and thus tends to overstate benefits[3]. It is thus reasonable to expect different central tendencies and variances among the different methods. Second, contingent valuation can measure nonuse values in addition to use values, whereas the weak complementarity methods cannot measure nonuse values. Thus, it is also reasonable to expect valuations from contingent valuation studies to be higher than similar valuations by the other methods[4]. Third, contingent valuation is an ex ante

[3] Desvousges, Johnson, and Banzhaf (1998) explain, "If there are substitutions for the site, an increase in travel costs would induce people to visit another site rather than forgo recreation altogether. This substitution means that not all the value of the trip is lost. It is only reduced by the difference in satisfaction given by the second-choice site relative to the first-choice site" (p. 20). The absence of substitution effects thus causes the method to overstate benefits.

[4] Freeman (1993) calls the verification of nonuse values as the most important question about them, rather than the motives for them. "Arguments about motivations seem to be offered primarily to persuade the reader of the plausibility of the hypothesis that nonuse values are positive. But the real test of this hypothesis will come from the data. Rather

approach to valuation while hedonic pricing and the travel cost method are ex post approaches. Ex ante analysis involves the prediction of the economic consequences of future or hypothetical events, while ex post analysis involves measuring the actual economic consequences of events. It is reasonable to expect ex post analyses to exhibit greater accuracy than ex ante analyses since measurement of past behavior is generally more accurate than prediction of future behavior[5]. Finally, as discussed in Chapter 1, there are theoretical reasons why contingent valuation results should be different than those from hedonic pricing and the travel cost method. This test of convergent validity will bring statistical evidence to all these explanations for divergence.

There are a few other variables that are less controversial but worth testing for their moderating effects. The value in testing them is more to control for their effects than for the insight to their effects themselves. Nonetheless, the degree of their effects can be instructive. The first variable is the incomes of respondents. Clearly the amount of income available to consumers limit and affect what they can consume (Ellis, 1998, p. 159; Freeman, 1993, p. 168). Second, the years of the data collection were tested (Ofiara & Seneca, 2001, p. 265). Values for the environment may change over time, so it is reasonable to expect environmental valuations to change with time as well. Finally, the

---

than further debating definitions and possible motivations, it would be more useful to proceed with a test of the hypothesis that nonuse values (defined in a way that makes testing of the hypothesis feasible) are positive" (p. 145). If the contingent valuation method is shown to be a moderating variable, then one explanation for it could be the verification of positive nonuse values. Furthermore, the magnitude of the nonuse values for specific goods could then be estimated by deducting revealed preference estimates for the goods from corresponding contingent valuation estimates (pp. 159-161).

[5] Freeman (1993) prescribes more use of ex post analysis to validate ex ante analyses. "It is particularly important that the economic analysis of environmental and resource policies include ex post analysis. It is necessary… to devote more effort to verifying these models through ex post comparisons of the predictions with observed results" (pp. 14-15).

locations of the studies were also tested. These results describe the geographic generalizability of valuation studies. For each of these variables, the null hypothesis is that it is not a moderating variable, and the alternative hypothesis is that it is.

Rival Hypotheses

It is likely, of course, that the environmental valuation data exhibit high variance even after accounting for the above potential moderators. I did not expect to identify sufficient moderators to explain all or even most of the variance. There are rival hypotheses beyond the scope of this dissertation that might further explain the variance. First, the variance might be due to some other variables not tested in this research. The variables tested in this dissertation are those whose effects, or the lack thereof, are instructive to the future use of the valuation methods and which I can measure in existing valuation studies. However, there could be other variables outside the measures of the studies that account for the dispersion of values.

Second, there may simply be a truly wide variance in the values of environmental goods. This is not unreasonable to expect because they are often nonmarket goods, and consumers are not accustomed to pricing them. Furthermore, even some marketed goods have wide ranging prices. For example, art work, which in some ways shares some of the aesthetic and intangible benefits of environmental goods, exhibit wide ranging values for similar goods. The wide ranges in values, then, could simply be a reflection of the diversity of views and values held for the environment.

Also, one of the valuation methods contributes to the variance in results. Randall, Hoehn, and Brookshire (1983) have argued that contingent valuation studies result in highly contingent responses. That is, responses to contingent valuation surveys are unique to the hypothetical markets described in the surveys. They vary with several aspects of the hypothetical market, including the amenity of interest, the status quo level of the amenity, the offered change in the level of the amenity, the institutional structure of provision, the method of payment, the decision rule to provide the change, etc. Thus, residual variance among contingent valuation results should be expected.

Data Collection

The above hypotheses were tested by statistical analysis of data collected from the existing body of environmental valuation studies. Following the meta-analytic framework, the research began with a thorough and systematic search for applicable studies, followed by the coding of information from the studies.

<u>Case, Unit of Analysis, and Sample Frame</u>

Meta-analysis is fundamentally a statistical analysis of the body of literature on a given topic. Thus, the basic artifact from which it gathers data, or its "case" (Vogt, 1999, p. 34), is the study. For this dissertation, the case is the environmental valuation study. The unit of analysis for this research is the valuation function. This could be an average willingness to pay, a trip generation function, a hedonic price function, or another quantitative expression of environmental value. Notice that more than one analytical unit can come from each case (Lipsey, 1994, pp. 112-114). For example, one study might use multiple methods to value a good, and thereby generate multiple valuation functions from one case. (However, when a study reports multiple valuation functions from the *same* data – such as when a study reports all the functional forms considered – only the most highly recommended one will be used.)

Meta-analyses of topics with small and known populations of studies often include the entire population in their analyses, thus reducing the analysis to descriptive statistics. It is not necessary, however, to include every paper on a topic in a meta-analysis. For well-published topics such as environmental valuation, it is not even desirable to do so. In such cases, a representative sample of studies is preferable, from which inferences on the population of studies can be made (White, 1994, p. 44). Thus, for this meta-analysis, I collected a sample of applicable studies representing the most current practices of the three environmental valuation methods. The specific parameters used to select studies for inclusion in the meta-analysis are as follows:

First, the studies were published in periodicals. This parameter is needed for a purely pragmatic reason: unpublished studies are largely unknown and inaccessible. They are not only inaccessible to a meta-analyst, but also to benefit transfer analysts searching for applicable studies. Because I am interested in the assessment of studies being used to value environmental goods, excluding unpublished studies is acceptable. There is also a compelling methodological reason for this parameter: publication is a signal of the quality of the study. Manuscripts submitted to academic journals typically undergo peer reviews. These reviews provide some level of assurance to the quality of the design, implementation, and results.

Second, the studies must employ the travel cost method, hedonic pricing, contingent valuation, or any combination of them. These three methods are at the heart of the research questions, so they are an essential parameter.

Third, the studies must quantitatively value a change in the quality of an environmental good or a change in the level of an environmental service. It is not enough to simply value a static good, such as the value of a stand of trees. Rather, the studies must value a change in the good, such as the value of preserving the stand of trees from being completely cut down. This condition is necessary to more clearly define the magnitudes of the environmental goods being valued. It is also necessary for the calculations of effect sizes, as discussed later in this chapter. Valuations of outdoor recreation were only included when they were related to changes in environmental quality. For example, the value of fishing in a stream would not be included in the meta-analysis, but the value of fishing in a stream before and after an increase in the stream's water quality would be included.

The search for studies meeting the above three parameters consisted of a manual search of key periodicals in environmental policy. This method is widely used by literature reviewers with favorable results (White, 1994, p. 45). Manual search is not an efficient method (that is, it captures many irrelevant articles per relevant article), but it virtually assures the researcher that all relevant studies in the frame of the manual search are found.

To determine the key journals to be manually searched, I searched five bibliographic databases pertaining to environmental issues. In each database I searched for journal articles pertaining to the travel cost method, hedonic pricing, and contingent valuation. The five databases included one general academic database (Academic Search Primer), two environmentally focused databases (Environmental Sciences and Pollution Management Set, and Aquatic Sciences and Abstracts & Oceanic Abstracts), an economic database (EconLit), and an agricultural database (AGRICOLA). For each database, key word searches used three phrases: "travel cost*," "hedonic," and "contingent valuation." The journals cited in these searches were recorded, as well as the number of hits for each journal. Altogether, 141 journals were cited, with 753 hits (including multiple hits from more than one database). However, over a third of all hits came from just five journals. (See Figure 3.1.) These top five journals, all available in local university libraries, were manually searched[6]:

- *American Journal of Agricultural Economics*
- *Environmental & Resource Economics*
- *Journal of Environmental Management*
- *Land Economics*
- *Water Resources Research*

To capture the most current practices used in each of the valuation methods, I started the literature search from December 2001 and worked back to the beginnings of the contemporary practices of each method, which provided a sufficient sample for the statistical analyses. The number of studies in a meta-analysis need not be large[7], but normal statistical criteria favor the use of as many observations as possible.

---

[6] The *Journal of Environmental Economics and Management* and *Ecological Economics* – two widely regarded journals in nonmarket valuation – ranked numbers 6 and 7 by the number of hits. Thus, they were not included in the sample frame for this research.
[7] Some meta-analyses in the medical field have had as few as three studies.

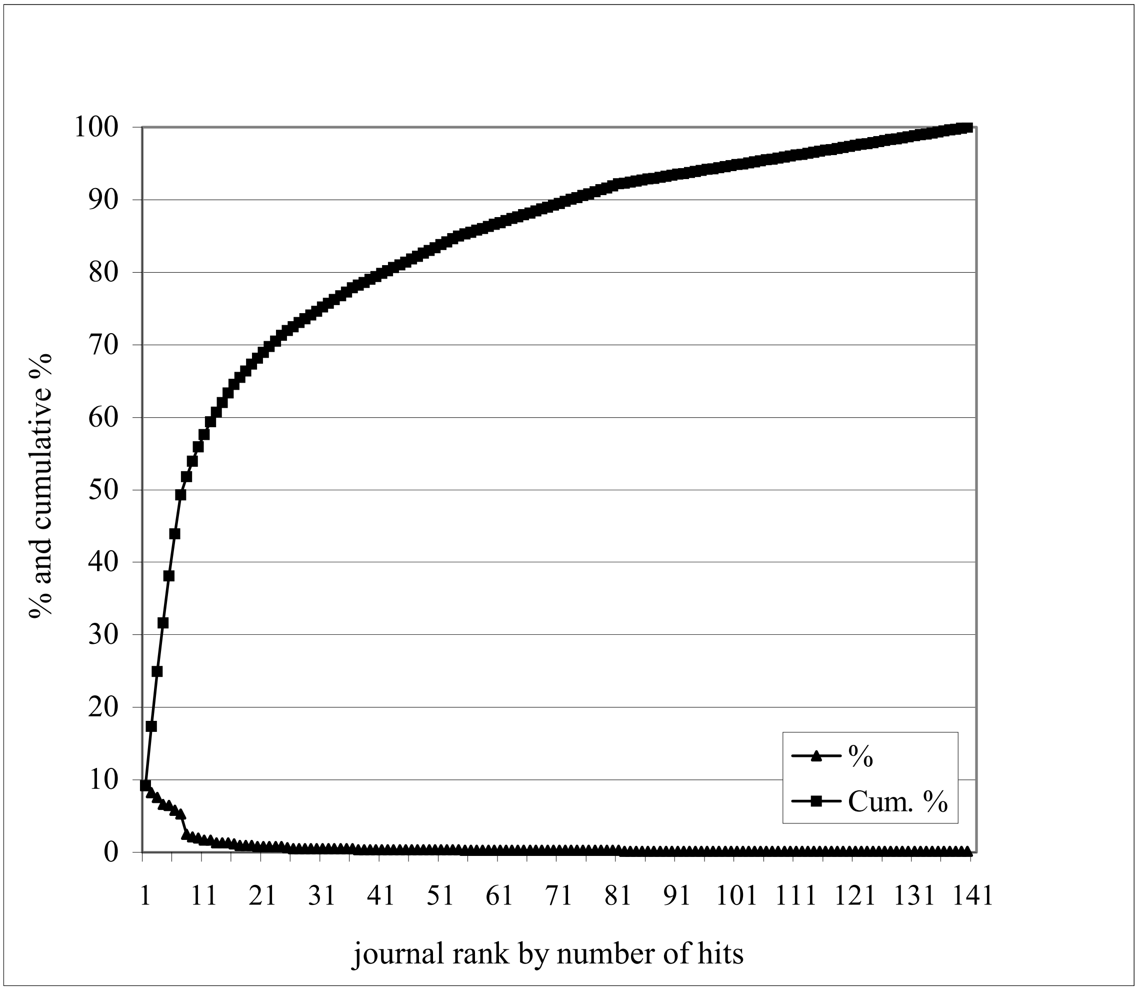


Figure 3.1: Percent Hits of Top Journals

Variables

From each case a list of variables was coded, and these variables provided the data for the statistical analyses. The information coded focused on those directly needed to test the research hypotheses (i.e., study results, method and substantive variables), but also included extrinsic variables such as journals and authors (Lipsey, 1994, pp. 112-116). Most of the information was directly transcribed from the studies and required little or no inference (e.g., sample size, response rate, etc.), a few required minor inference (e.g., type of effect, effect size, etc.), and none required high inference (e.g., quality of the study) (Hall, Tickle-Degnen, Rosenthal, & Mosteller, 1994, pp. 25-26).

Foremost in importance was the recording of each study's results. This not only included the specific quantification of value (e.g., average willingness to pay, hedonic price function, trip generation function), but also reported statistics related to the value, such as variance (Mitchell & Carson, 1989, pp. 307-308). Another critical variable is the method of valuation (Ofiara & Seneca, 2001, p. 265; Stock, 1994, p. 128). At the surface level this includes the broad type of valuation tool used (travel cost method, hedonic pricing, or contingent valuation), and whether multiple methods were used in the study (Mitchell & Carson, 1989, pp. 307-308). For each method, however, additional details were also coded to further qualify it. The variables coded for each method included the following[8]:

Travel Cost Method

- Variation of the method (Hanley & Spash, 1993, p. 91; Hellerstein, 1995): individual or zonal origins of travel.
- Type of dependent variable (Freeman, 1993, p. 444): number of trips to a site, whether to visit a site.

[8] Literature indicating the potential influence (i.e., moderating effect) of certain variables in the outcomes of the methods are cited.

- Medium of the survey: telephone, mail, in-person, focus group, or others.

Hedonic Pricing (Freeman, 1993, pp. 375-379)

- Weakly complementary good: residential property, agricultural property, commercial property, or others.
- Source of the dependent variable: actual market transactions, census tract data on house values, professional appraisals, or others.
- Functional form of the dependent and environmental variables (Chattopadhyay, 1999).
- Explanatory variables: the numbers of explanatory variables describing sites, neighborhoods, environmental goods, and others.

Contingent Valuation (Mitchell & Carson, 1989, pp. 301-308)

- Elicitation method: iterative bidding or bidding game, direct or open-ended question, payment card or discrete choice, referendum with or without follow-up, or others. (Alberini, 1995; Balistreri, McClelland, Poe, & Schulze, 2001; Boyle et al., 1996; Brown, Champ, Bishop, & McCollum, 1996; Halvorsen, 1998; Huang & Smith, 1998; Jordan & Elnagheeg, 1994; Mattsson & Li, 1994; Morrison, 2000; Ready, 1996; Ready, Navrud, & Dubourg, 2001; Reaves, Kramer, & Holmes, 1999; Scarpa & Bateman, 2000; Whitehead, Blomquist, Ready, & Huang 1998)
- Payment vehicle: taxes, fees, price increase, etc.
- Type of dependent variable: willingness to pay or willingness to accept. (Adamowicz, Bhardwaj, & Macnab, 1993; Knetsch, 1990; Kolstad & Guzman, 1999; Loehman, Park, & Boldt, 1994; Mansfield, 1999)
- Medium of the survey: telephone, mail, in-person, focus group, or others. (Ethier, Poe, Schulze, & Clark, 2000; Loomis & King, 1994; Mannesto & Loomis, 1991)
- Separation of estimates: use and nonuse, or combined. (Bishop, Champ, Brown, & McCollum, 1997; Butte & van Kooten, 1999)

Next, I coded substantive variables central to the hypotheses presented. These included the type of environmental good, the magnitude of the environmental change, the type of description used, and demographic information:

Type of good

- The affected services or commodities such as air, water, fishery, forest, etc. (Freeman, 1993, pp. 12-14; Ofiara & Seneca, 2001, p. 265; Smith, 1995)

Magnitude of the environmental change

- The baseline and ending environmental quality, and the change (Michael, Boyle, & Bouchard, 2000).
- Temporal and spatial boundaries of the change.
- Type of effect (Freeman, 1993, pp. 12-14) or treatment (Stock, 1994, p. 128), such as direct impacts on humans (morbidity, mortality, visibility, odor, visual aesthetics), ecosystem effects (impacts on productivity of ecological systems [forestry, fisheries], recreational uses, biodiversity, stability), impacts through nonliving systems (weather, climate, materials, soils), etc.

Type of description of environmental change (Boyle, 1989; Michael, Boyle, & Bouchard, 2000)

- Quantitative versus qualitative descriptions of change. That is, objective measures that are reproducible (e.g., dissolved oxygen in water) or qualitative perceptions of change.
- Incremental versus holistic change.
- Numbers of measures used (Freeman, 1993, pp. 469-461), such as one measure, one summary measure or index, or multiple measures.

Other substantive variables

- Income of the study's unit of analysis.
- Year of data (Mitchell & Carson, 1989, pp. 301-303; Ofiara & Seneca, 2001, p. 265; Whitehead & Hoban, 1999).
- Location of the study and its setting (Stock, 1994, p. 127), such as local, regional, national, international, etc.
- Descriptions of the subjects, including both the sample frame and the population it is meant to represent; the original sample size, response rate, and useable number of responses; and whether probability or non-probability sampling was used (Dalecki, Whitehead, & Blomquist, 1993; Mitchell & Carson, 1989, pp. 301-303; Stock, 1994, pp. 127-128).

Finally, some extrinsic variables (Lipsey, 1994, p. 114) were also coded, including bibliographic study identifiers (Stock, 1994, p. 127). Any findings and comments relevant to this research were also recorded. The code sheet used for this data collection is in Appendix A.

Analysis

Analysis of the resulting data consisted of descriptive statistics, calculations of effect sizes, and tests for moderating variables. First, descriptive statistics of the data summarized the distribution of valuation studies by good, valuation method, year, and location. The measures of value varied, including willingness to pay, willingness to accept, trip generation functions, and hedonic price models. Thus, standardized effect sizes were calculated for each environmental valuation, and these effect sizes were summarized by descriptive statistics.

Effect sizes are standardized measures of the impact the independent variable has on the dependent variable. More specifically, it is "A statistic… indicating the difference in outcome for the average subject who received a treatment from the average subject who did not" (Vogt, 1999, p. 94). This definition stems from the medical sciences, which uses meta-analysis extensively. In a medical context, for example, an effect size might measure the difference in the average recovery time between patients who received a treatment and those who did not. In this current research, an effect size is the difference in value between an environmental good that is subject to some degradation or improvement and one that is not.

Calculations of effect sizes vary according to the information reported in the original studies, but follow the general form

$$d = (Y_E - Y_C)/S,$$

where Y_E - Y_C is the difference in average effect (valuation) before and after the treatment (change in environmental quality), and S is the pooled standard deviation from both cases (Hunter & Schmidt, 1990, pp. 233-235). The effect size is unitless, thus facilitating comparability between studies.

Methods of calculating effect sizes from various statistics and models are described in several sources (Fleiss, 1994; Hunter & Schmidt, 1990; Rosenthal, 1994; Rosenthal & Rosnow, 1991; van den Bergh, Button, Nijkamp, & Pepping, 1997), but patterns did emerge in this data set. When studies used the travel cost method to value a change in an environmental good, they typically did so by some form of a two-sample test. In one procedure, visitation behaviors are examined before and after some change in the environmental quality of a site (e.g., change in visits to a lake before and after water quality improvements). This is essentially a pre-post test, and the effect size, d, was directly calculated with Y_E - Y_C representing the change in visitation before and after the environmental quality change. In another procedure, two sites similar by several characteristics except the environmental quality in question are compared (e.g., difference in visitation between a clean lake and a polluted one). Again, in these cases effect size were directly calculated with Y_E - Y_C representing the difference in visitation between two comparable sites. Sometimes this latter procedure was expanded to more than two sites with varying levels of the environmental good in question. In this cross-sectional approach, the level of the environmental good becomes an independent variable in the trip generation function, and the coefficient of that variable is the change in visitation for a unit change (for a ratio level measurement) or a discrete change (for a nominal measurement) in the environmental good, Y_E - Y_C .

When studies used hedonic pricing to value an environmental good, they almost always followed this last example of the travel cost method: the level of the environmental good was used as an independent variable in the hedonic pricing function. In these cases, the implicit price[9], r, of the environmental good is the difference in price at two levels of the environmental good. That is, $Y_E - Y_C = r = \delta P/\delta Q$. Again, when the environmental good is measured with ratio level data, then the price is for an increment of environmental change. When it is measured with nominal data, then the price is for a discrete or holistic environmental change.

Finally, when studies used contingent valuation, the reported mean willingness to pay (or willingness to accept) represents the change in value for the change in the environmental good. That is, the mean WTP = Y_E - Y_C . This assumes that the mean WTP for the status quo environmental condition (control situation), Y_C , is zero, while the mean WTP for the improved condition (experimental situation), Y_E, is the stated WTP. This is a reasonable assumption because contingent valuations ask respondents their WTP *above current expenditures* for the improved environmental condition.

With effect sizes calculated, the heart of the meta-analysis proceeded with the combining of effect sizes while adjusting for sample size, and testing the presence of moderators with tests of homogeneity (van den Bergh, Button, Nijkamp, & Pepping, 1997, pp. 70-72). Homogeneity tests check the assumption that there exists a common

[9] See Chapter 1.

effect size among studies. Rejecting this assumption leads to the search for variables explaining the variance.

The methods of testing specific potential moderating variables vary with the characteristics of the data available. Meta-analyses have used non-parametric statistics, statistical correlation, differences in means tests, and regression analysis to test hypotheses (van den Bergh, Button, Nijkamp, & Pepping, 1997, p. 51). In this research, moderating effects were tested primarily through the difference in means tests (e.g., t-tests and analyses of variance) and regression analyses.

Troubleshooting

The elegant simplicity of the meta-analytic framework hides the many difficulties meta-analysts often face. A small-scale, pilot meta-analysis that I prepared under a grant from the U.S. Environmental Protection Agency, and presented to the National Center for Environmental Economics in March 2001, revealed problems that I expected to face in this dissertation research[10]. It helped me to address those issues as I conducted this comprehensive meta-analysis. One of the anticipated problems was defining a reasonable population frame for this analysis. The population of environmental valuation studies is too large to include them all in this meta-analysis. Thus, I carefully defined a limited sampling frame of studies, described earlier in this chapter. Another anticipated problem was the poor reporting standards in environmental valuation studies. Carson, Flores,

[10] The purpose of the pilot study was specifically to test the feasibility of conducting an extensive meta-analysis of environmental valuation as described here. It did so by conducting a mini-meta-analysis of the value of clean water. The results of the pilot study were positive: not only is it possible to conduct a large-scale meta-analysis on environmental valuation, but the results identified significant moderators that need verification in this dissertation.

Martin, and Wright (1996) described this problem in the aftermath of their meta-analysis, saying, “Our efforts to conduct this analysis… have been greatly hindered by the curse suffered by other meta-analyses of nonmarket data: incomplete reporting of necessary details” (p. 98). In the pilot meta-analysis I performed, 46% of applicable studies provided sufficient information to be included in the meta-analysis. Thus, I anticipated having to collect about twice the number of studies as is ultimately used in this meta-analysis.

CHAPTER 4

DATA

This chapter summarizes and describes the data collected for this research. It begins with an accounting of the cases and units of analysis gathered. Then, through a series of graphs and tables, it describes the distributions of the major variables measured, their relationships to each other, and temporal trends in the data. These findings constitute the first anticipated result discussed in Chapter 2, descriptive statistics of existing environmental valuation studies.

Cases and Units of Analysis

The five key journals were manually searched. I examined the titles, abstracts, and contents of each article in each issue of these journals from 1970 to 2001. Two of the five journals began their publication during this period. The *Journal of Environmental Management* published its first volume in 1973, and *Environmental and Resource Economics* published its first volume 1991. The other three were well established by 1970. Of all these journal issues, only a few were not available in any local library.[1, 2]

The manual search resulted in over 400 articles dealing with some aspect of these environmental valuation methods. These articles were collected for further assessment and coding. This resulted in a total of 228 articles (cases) that met all the selection

---

[1] Journal holdings of Atlanta area libraries were searched in the Georgia Libraries Journal List (GOLD) database.
[2] All issues of *Environmental and Resource Economics* in 1991 and the last four issues in 2001 (September through December) were not available at the time of data collection.

criteria discussed in Chapter 3. (See Appendix B for a complete listing of these articles.) These articles were analyzed in depth and their information was coded. The 228 cases resulted in a total of 614 valuations of changes in environmental goods (units of analysis). Not all of these valuations contained information on all the variables I wished to code, but each had information on at least some of the variables. Therefore, different subsets of this total number were used for different data analyses in this chapter and the next.

Descriptive Statistics of the Primary Independent Variables

The 614 valuations are described below by their distributions over the four primary independent variables of interest: environmental good, magnitude of change, description of change, and method of valuation. Descriptive statistics of the dependent variable, effect sizes, are presented in Chapter 5.

<u>Environmental Good</u>

The specific good valued by each unit of analysis was recorded as described in each study. These goods varied considerably by specifics such as location and measures of quality. To summarize characteristics of the data, they were aggregated under broader categories. For example, valuations of clarity in Lake Tahoe and body-contact safety of coastal waters were combined under a category for "water quality", while valuations of lake levels in Mono Lake and flows in the Wisconsin River were combined under a category for "water quantity". Of course, these constructed categories were not always definitive. The level (water quantity) of Mono Lake, for example, affects its salinity (water quality) and its suitability for wildlife. In such cases, the good category was determined by what subjects were actually asked to value or assumed to have valued through their market behavior. Still, judgment was used in these determinations. To measure even broader characteristics of the data, these two water categories were further combined into the general category of good called "water". Similar categories and subcategories were developed for most goods valued. The distribution of goods valued in the data set is summarized in Figure 4.1.

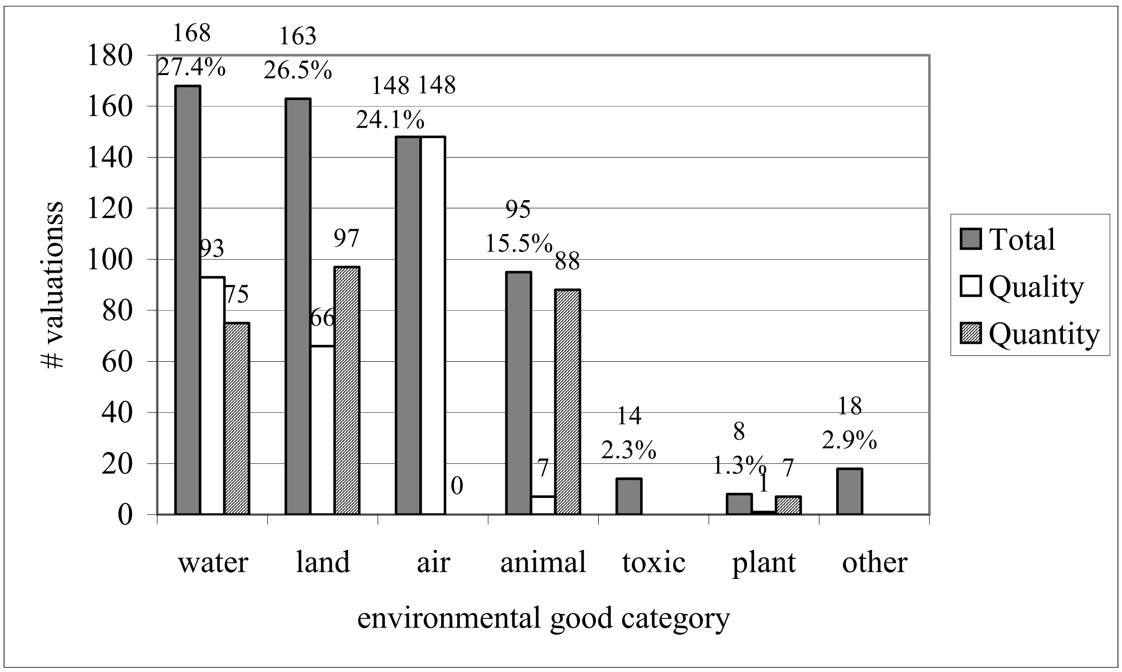


Figure 4.1: Distribution of Valuations by Environmental Good

Water, land, and air are clearly the most often valued goods in this data set, with each representing about a quarter of the observations. Animals are fourth, accounting for 15.5% of the valuations. Several goods follow with relatively small numbers: toxics (14), plants (8), noise (4), waste (3), biodiversity (2), energy (2), general environment (2), recycling (1), and combinations of goods (4).

Like water, land goods were subcategorized as land quality and land quantity. The former refers to characteristics of a given area of land, such as soil quality, elevation, ground cover, etc. The latter refers to the amount of a given type of land, such as the acreage of designated wilderness. Animal and plant valuations were also divided into subcategories of quality and quantity. The quality of an animal or plant includes its health, size, etc. The quantity of animals or plants refers to their population in an area. This includes the popular valuation of saving a species from extinction.

The air category only had valuations of its quality. This is not surprising since humans have thus far been more concerned with, and affected by, the quality of ambient air than its quantity. Also, human-induced activities such as combustion and deforestation have had a more significant impact air's quality than quantity.

Each of these categories of goods includes many specific goods. Examples of these are given in Table 4.1. The table reveals an important point that must be made about these categories and their use in the meta-analysis: all but one of the categories have multiple goods that we can expect to account for much of the variance in valuations.[3] The animal category, for example, includes species ranging from the common black fly to the exotic bighorn sheep and the endangered spotted owl. One might reasonably expect the human valuations of these diverse species to also be diverse. However, the air category only has one good: air quality. It does not have multiple goods

[3] Unfortunately, there are insufficient numbers of valuations for each of these goods to test this claim statistically. Thus, it was necessary to aggregate the data into the broader categories of goods.

Table 4.1: Constituents of Environmental Goods

Good	Example constituents
Water	aquatic system, bay, drinking water, estuary, flooding, groundwater, harbor, irrigation, lagoon, lake, natural water, ocean, pond, rainfall, river, sea, sound, stream, surface water, swimable water, water basin, water front, water quality, water supply
Land	beach, countryside, creek side, development, elevation, forest, grassland, green space, habitat, landfill, marsh, natural land, open space, park, rain forest, shoreline, soil, Superfund site, wetland, wilderness, woodland
Air	air quality
Animal	bear, bighorn sheep, black fly, caribou, coyote, deer, eagle, elk, fish, game, moose, pheasant, possum, salmon, seal, spotted owl, squawfish, striped shiner, trout, walleye, whale, whooping crane, wild turkey, wildlife, wolf, woodpecker
Toxic	hazardous waste, heavy metals, lead, pesticides
Plant	aquatic plants, daffodils, trees, wildflowers
Other	biodiversity, combinations of goods, energy, environment, noise, recycling, waste

to explain some of the variance. Thus, when testing the effects of the independent variables on valuation, as discussed in the next chapter, I expected the subset of air studies to have more of its variance explained.

Magnitude and Description of Change

Two other primary independent variables are the magnitudes of the changes in the environmental goods, and the descriptions of the changes. As discussed in Chapter 3, these variables were also categorized in order to characterize the data set, and to aggregate the data for statistical analysis.

The magnitudes of the changes in the environmental goods were categorized as incremental and holistic changes. For example, a unit change in tropospheric ozone concentrations would be categorized as an incremental change of air quality, while a change in ozone concentrations to meet air quality standards would be categorized as a holistic change. Similarly, a unit change in acreage of woodlands would categorized as an incremental change, while an addition of an entire park would be categorized as a holistic change. These categories are admittedly arbitrary (e.g., the addition of one park could be incremental when compared to the addition of several parks), and they reduce all levels of changes to just two categories, but the data reduction was desirable for several reasons. First, the variance in the changes bordered on each being unique. Indeed, there are few studies in the data set that valued the same changes in an environmental good. Thus, data reduction was necessary to develop sufficient degrees of freedom for statistical analysis. Second, the uniqueness of the changes is a problem also faced by practitioners of benefit transfer. To overcome this problem, they seek studies that are similar in environmental changes, not identical, to the current situation they study. Thus, the categorization of changes help policy analysts by identifying the relevant types of changes to consider when reviewing existing valuation studies. Third, while the categories are arbitrary to some extent, they are still meaningful. They represent valuations at the margin (incremental changes) versus those for outcome-oriented, new states of quality or quantity (holistic changes). This difference is hypothesized to moderate valuations. Fourth, this hypothesis can be tested with the reduced data.

The descriptions of the changes were categorized as qualitative and quantitative. For example, a 10% reduction in biochemical oxygen demand in a river would be categorized as a quantitative description, while an improvement of water quality to "clean" or "safe" levels would be categorized as a qualitative description. The justification for this data reduction is similar to that for the magnitude of changes: the high variance in the data, bordering on uniqueness, necessitated data reduction for statistical purposes; and the insights that could be gained from data reduction are desirable.

Tables 4.2 and 4.3 display the frequency distributions of these two variables, and Table 4.4 displays their cross-tabulation. Table 4.2 shows that both categories of changes are well represented in the data set, with 58.3% of changes being holistic and 41.0% being incremental. Similarly, Table 4.3 shows that both types of descriptions are often used, with 57.2% being quantitative and 42.5% being qualitative. However, the cross-tabulation of these two variables show a strong relationship between them. Specifically, qualitative descriptions tend to be associated with holistic changes, and quantitative

Table 4.2: Distribution of Valuations by Magnitude of Change

Change	Frequency	% Frequency
Holistic	358	58.3
Incremental	252	41.0
Unknown	4	0.7
Total	614	100.0

Table 4.3: Distribution of Valuations by Description of Change

Description	Frequency	% Frequency
Qualitative	261	42.5
Quantitative	351	57.2
Unknown	2	0.3
Total	614	100.0

Table 4.4: Magnitudes of Change Used with Descriptions of Change

	Qualitative	**Quantitative**	Total
Holistic	234 (38.4%)	124 (20.3%)	358
Incremental	27 (4.4%)	225 (36.9%)	252
Total	261	349	610

$\chi^2 = 180.4$

descriptions tend to be associated with incremental changes. Conversely, qualitative descriptions tend not to be used with incremental changes, and quantitative descriptions are less frequently used with holistic changes. The χ^2 (chi square) test of independence[4] for categorical data resulted in a test statistic of 180.4, statistically confirming the relationship at a level of significance of $\alpha = 0.01$.[5] The relationship is not surprising. Quantitative measures of changes are ratio or interval types of data, and are thus amenable to incremental aggregation or disaggregation. Air quality, for example, can be quantitatively measured by concentrations of particulates or other pollutants, so its valuation can be made in increments of changes. Qualitative measures, on the other hand, are ordinal or categorical types of data, and are not reliably amenable to aggregation. So a description of air quality change from poor to excellent would tend to be valued holistically. There are exceptions to both these relationships, as indicated in Table 4.4. Sometimes, quantitative measures are used in holistic valuations, such as when air quality is measured in concentrations of pollutants, but the change valued is that to meet air quality standards. In such a case, the change is not a unit of concentration but a jump to safe air quality levels. Other times, ordinal qualitative measures are used in incremental valuations, such as when subjects are asked to value changes in air quality

---

[4] The χ^2 test of independence measures the relationship between two categorical variables by comparing observed cross-tabulation distributions with expected distributions if the variables are unrelated (i.e., independent). Formally, the null and alternative hypotheses for the χ^2 test of independence are

H_0: the variables are independent

H_1: the variables are not independent

The null hypothesis is rejected if the test statistic is greater than the critical value, defined by the degrees of freedom (d.f.) in the table, and a selected level of significance, α. Otherwise the null hypothesis cannot be rejected.

[5] The critical value of χ^2 for 1 d.f. and $\alpha = 0.01$ is 6.6. The test statistic is greater than the critical value, therefore the null hypothesis of independence is rejected.

between fair, poor, good, and excellent. Table 4.4 shows, however, that this is rarely done.

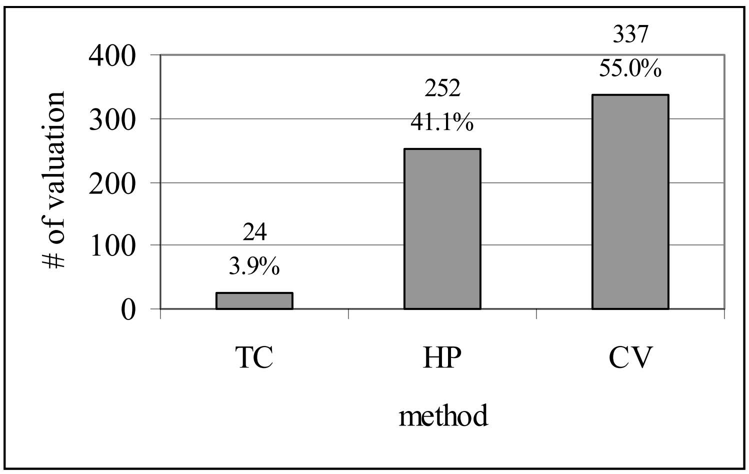
Valuation Method

The primary independent variable of interest is the method of valuation. Figure 4.2 illustrates the distribution of observations by the valuation methods used. Included in these numbers are only those studies that applied one or more of these three methods to value environmental goods. Excluded from these numbers are the many studies in these journals that use these methods to value non-environmental goods[6]. For example, many studies used the travel cost method to value recreational goods such as golfing, hunting, fishing, rafting, etc. Many others used hedonic pricing to measure non-environmental determinants of property value, such as racial make-up of neighborhoods, zoning, electric power lines, etc. Still others used contingent valuation to measure market goods such as community infrastructure, safety measures, licenses and permits, etc.

The contingent valuation method (CV) accounts for a majority of the valuations, at 55.0%. This is somewhat surprising because it is the youngest of the three methods. However, it is also the most versatile of the three, prescribed for all sorts of environmental goods when the other two are more limited in their applicability (Freeman, 1993, pp. 486-487). The figure also shows that the travel cost method (TC) is rarely used to value changes in environmental goods, accounting for just 3.9% of the valuations. Many more travel cost studies than 24 were found during the manual search of the five

[6] While these methods were developed in part to address the gap between environmental policy analysis and cost-benefit analysis (see Chapter 1), they have broader applications in valuing non-environmental and market goods.

key journals, but the vast majority of them were valuations of outdoor recreation, not changes in environmental goods. Indeed, the adaptations necessary to use the travel cost method to value changes in environmental goods are significant and make alternative methods more appealing. Finally, the hedonic pricing method (HP) stands in the middle, accounting for 41.1% of the units of analysis.



TC = travel cost, HP = hedonic pricing, CV = contingent valuation

Figure 4.2: Distribution of Valuations by Valuation Method

Relationships Among the Variables

Besides the relationship between the magnitudes and descriptions of environmental changes, other relationships between the primary independent variables are also noteworthy. The cross-tabulations and χ^2 test statistics for the relationships between method of valuation and the other independent variables are presented in Tables 4.5 through 4.8. Table 4.5 displays the frequency distributions of the methods used for each major category of environmental goods. The χ^2 statistic for this contingency table is 132.1, far greater than the critical value of 26.2 for 12 d.f. and $\alpha = 0.01$. However,

Table 4.5: Methods Used to Value Goods

	TC	HP	CV	Total
Water	15 (2.5%)	74 (12.1%)	78 (12.7%)	167
Land	1 (0.2%)	73 (11.9%)	89 (14.5%)	163
Air	0 (0.0%)	97 (15.8%)	51 (8.3%)	148
Animal	7 (1.1%)	1 (0.2%)	86 (14.1%)	94
Toxic	1 (0.2%)	1 (0.2%)	12 (2.0%)	14
Plant	0 (0.0%)	1 (0.2%)	7 (1.1%)	8
Other	0 (0.0%)	5 (0.8%)	13 (2.1%)	18
Total	24	252	336	612

Table 4.6: Methods Used to Value Goods – Reduced Table

	WC	CV	Total
Water	89	78	167
Land	74	89	163
Air	97	51	148
Animal	8	86	94
Other	8	32	40
Total	276	336	612

WC = weak complementarity method
$\chi^2 = 90.5$

several cells in the table have very few observations which may skew the test results.[7] Thus, the relationship between method and good was retested with the reduced cross-tabulation in Table 4.6. In this table, the travel cost and hedonic pricing methods were combined into a "weak complementarity" (WC) category, reflecting their common approach of indirectly measuring values through market behaviors for associated goods, and contrasting them against contingent valuation's direct valuation through hypothetical market behaviors. Also, the toxic and plant categories were combined with the "other" category. The reduced table confirms the relationship between method of valuation and environmental good. The χ^2 statistic is 90.5, greater than the critical value of 13.3 for 4 d.f. and $\alpha = 0.01$. Comparison of the observed distribution with the expected distribution under the null hypothesis[8] reveals where the relationships are strong. Animals are almost exclusively valued with contingent valuation, and air has a significant tendency to be valued with hedonic pricing. Water, land, and other goods exhibit no significant relationship with the method of valuation.

Table 4.7 shows a very strong relationship between method of valuation and the magnitude of change valued. The χ^2 test statistic is 326.8, compared to its critical value of 9.2 for 2 d.f. and $\alpha = 0.01$. Specifically, holistic changes tend to be valued with contingent valuation, while incremental changes tend to be valued with hedonic pricing. This is expected. Few contingent valuation studies ask respondents to value increments of changes. Instead, contingent valuation is quite apt to value end-state outcomes as they

---

[7] A common heuristic for the χ^2 test of independence requires at least 5 observations in each cell.

[8] The expected distribution is not shown but is easily calculated as (r x c)/t for each cell, where r, c, and t are the row total, column total, and total in the observed distribution, respectively.

may be easier to understand by respondents than incremental changes. The strength of hedonic pricing, on the other hand, is its ability to handle quantitative, ratio data. It can thus measure incremental changes easily. It is occasionally used, however, to measure holistic changes through bivariate independent variables representing the two alternative conditions of environmental goods. Indeed, 44 valuations in the data set did so.

Table 4.7: Methods Used with Magnitudes of Change

	TC	**HP**	**CV**	Total
Holistic	7 (1.1%)	44 (7.2%)	306 (50.2%)	357
Incremental	16 (2.6%)	206 (33.8%)	30 (4.9%)	252
Total	23	250	336	609

$\chi^2 = 326.8$

Table 4.8: Methods Used with Descriptions of Change

	TC	**HP**	**CV**	Total
Qualitative	7 (1.1%)	45 (7.4%)	208 (34.0%)	260
Quantitative	16 (2.6%)	207 (33.9%)	128 (20.9%)	351
Total	23	252	336	611

$\chi^2 = 115.7$

Similarly, Table 4.8 shows a strong relationship between method of valuation and the type of description. Contingent valuation tends to use qualitative descriptions of the changes in environmental goods, while hedonic pricing tends to use quantitative descriptions. This too is not surprising given the strong relationship between the magnitude and description of changes, described above.

Summarizing the uses of the valuation methods, the data provides statistical evidence that contingent valuation is associated with valuations of qualitative, holistic changes in all sorts of goods, but especially animals. Hedonic pricing is associated with valuations of quantitative, incremental changes in water, land, and especially air. The travel cost method is rarely used to value changes in environmental goods. Instead, it is more commonly used to value outdoor recreation.

Turning the focus to the environmental goods, the relationship between the environmental good and the magnitude of change is given in Table 4.9, while that between the good and its description of change is given in Table 4.10. Table 4.9 shows a strong relationship between the types of changes and the types of environmental goods.

Table 4.9: Magnitudes of Change Used with Goods

	Water	**Land**	**Air**	**Animal**	**Other**	Total
Holistic	89 (14.6%)	105 (17.2%)	41 (6.7%)	89 (14.6%)	34 (5.6%)	358
Incremental	78 (12.8%)	56 (9.2%)	107 (17.5%)	6 (1.0%)	5 (0.8%)	252
Total	167	161	148	95	39	610

$\chi^2 = 124.5$

Table 4.10: Descriptions of Change Used with Goods

	Water	**Land**	**Air**	**Animal**	**Other**	Total
Qualitative	73 (11.9%)	86 (14.1%)	40 (6.5%)	44 (7.2%)	12 (2.0%)	255
Quantitative	94 (15.4%)	77 (12.6%)	108 (17.6%)	51 (8.3%)	27 (4.4%)	357
Total	167	163	148	95	39	612

$\chi^2 = 24.3$

The χ^2 test statistic is 124.5, greater than the critical value of 13.3 for 4 d.f. and $\alpha = 0.01$. Comparison of the observed distribution with the expected distribution under independence reveals that the relationship is strongest with air and animals. Valuations of air tend to measure incremental changes. This reflects the common use of hedonic pricing to value changes in air quality. In these studies, air quality is usually measured as concentrations of specific pollutants. This ratio type data in the hedonic models result in valuations of incremental changes. Valuations of animals tend to measure holistic changes. This reflects a common scenario used in contingent valuations of animals: respondents are asked for their willingness to pay for preventing the extinction of an endangered species or to increase its population to some viable number. They are not typically asked to value an individual animal. Land and water, on the other hand, are not strongly associated with either type of change.

Table 4.10 shows only a very mild relationship between type of description and the environmental goods. The χ^2 test statistic is 24.3, just greater than the critical value of 13.3 for 4 d.f. and $\alpha = 0.01$. Changes in air quality are mildly more likely to be described quantitatively. This again reflects the popular use of hedonic pricing to value air quality. Also, changes in land are mildly more likely to be described qualitatively. Animals, water, and other types of goods do not show significant association with the type of description. This weak relationship between the description of change and the type of good is somewhat surprising. We have already seen that there is a strong relationship between the magnitude of change and the description of change, and we have seen that the magnitude of change is strongly associated with the type of good. Thus it is reasonable to expect the description change to be strongly associated with the type of good. Instead, the association is weak. This shows that the magnitude of change and the description of change have different effects on the other variables, so both should be included in the meta-analyses and regression analyses in Chapter 5. That is, the two variables are indeed measuring different things.

Other Variables of Interest

Besides these four primary independent variables, there are others of interest. Distributions of valuations by journal and region describe the specialties of the journals, and identify the regions that have received much attention as well as those that have received little attention. First, the frequency distributions of the two variables are presented in Figures 4.3 through 4.5. *Land Economics* (LE) by itself accounts for a majority of the valuations, even though it accounts for only a plurality of the cases. This is due to the many cases in that journal that report more than one valuation. The *Journal of Environmental Management* (JEM) and *Environmental and Resource Economics* (ERE) are next, accounting for 13.4% and 8.8% respectively, even though they are the youngest of the five journals. This reflects a focus of these two journals on environmental valuation. The *American Journal of Agricultural Economics* (AJAE) and *Water Resources Research* (WRR) are last, accounting for 8.5% and 6.2% of the valuations, respectively.

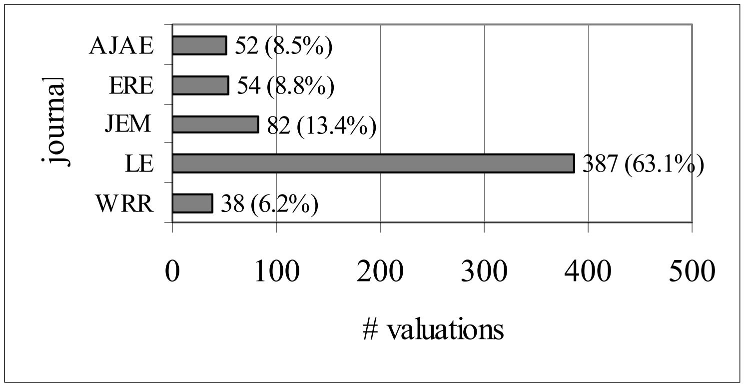


Figure 4.3: Distribution of Valuations by Journal

Figure 4.4 illustrates the distribution of valuations by country. This indicates the locations of the subjects valuing environmental goods, not the necessarily location of the goods. As it shows, 83.2% of the valuations in the data set were done in the United States.[9] Certainly a portion of this lopsidedness is due to the language of the search and articles for this data collection. The search was done with English language bibliographic databases and resulted in English language journals and articles. These studies, in turn, presumably tended to be done in English language countries. Notice that the leading four countries – United States, United Kingdom, Australia, and Canada – all use English as a

[9] It is important to note that this does not necessarily represent the distribution of the population of valuations worldwide. Recall that these five journals were the top cited journals in bibliographic searches on environmental valuation, not a representative sample of them.

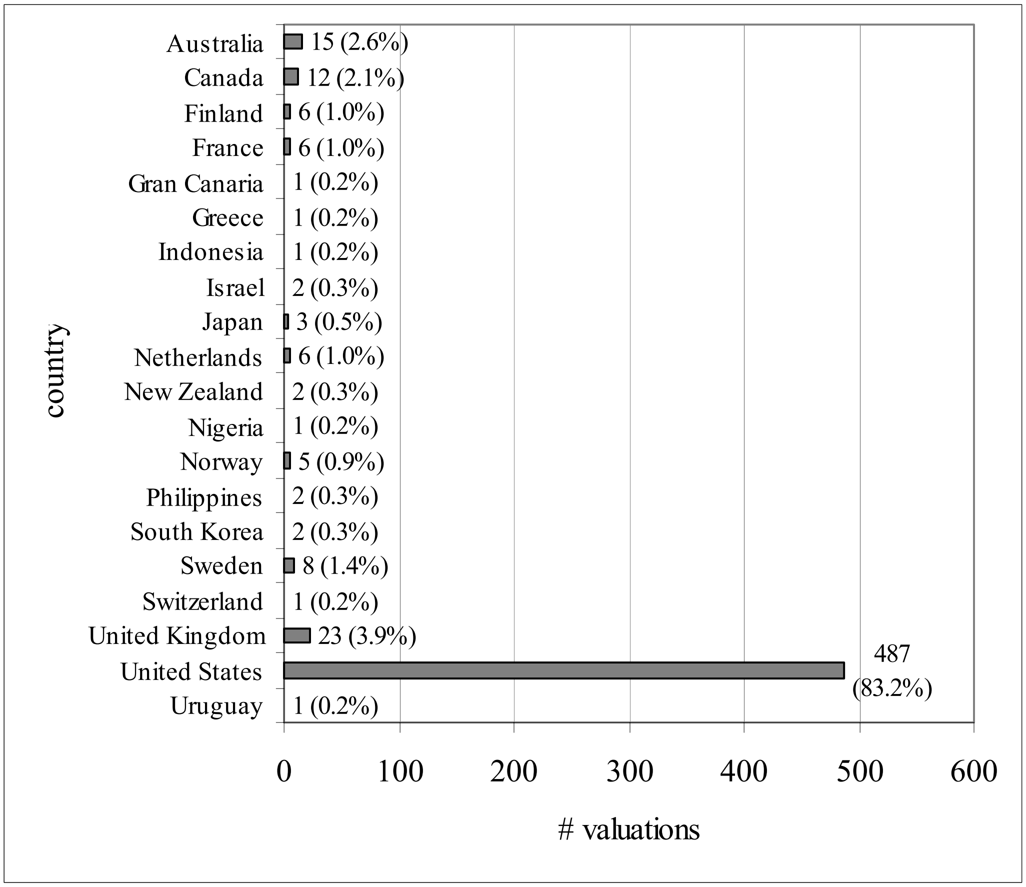


Figure 4.4: Distribution of Valuations by Country

primary language. Still, the dominance of the United States, relative to even the other English language countries, is due in part to the history of environmental valuation there. As discussed in Chapter 2, government prescription of cost-benefit analysis and environmental valuation tools has been strong in the U.S.

Of the data set's valuations done in the United States, Figure 4.5 shows their distribution by state. Western states appear to be well represented, while there is a noticeable lack of studies in Southern states. A cursory evaluation of the authors of the articles suggests that the geographic distribution is influenced by the locations of the authors' institutions. That is, it appears that authors tend to study local valuations. For

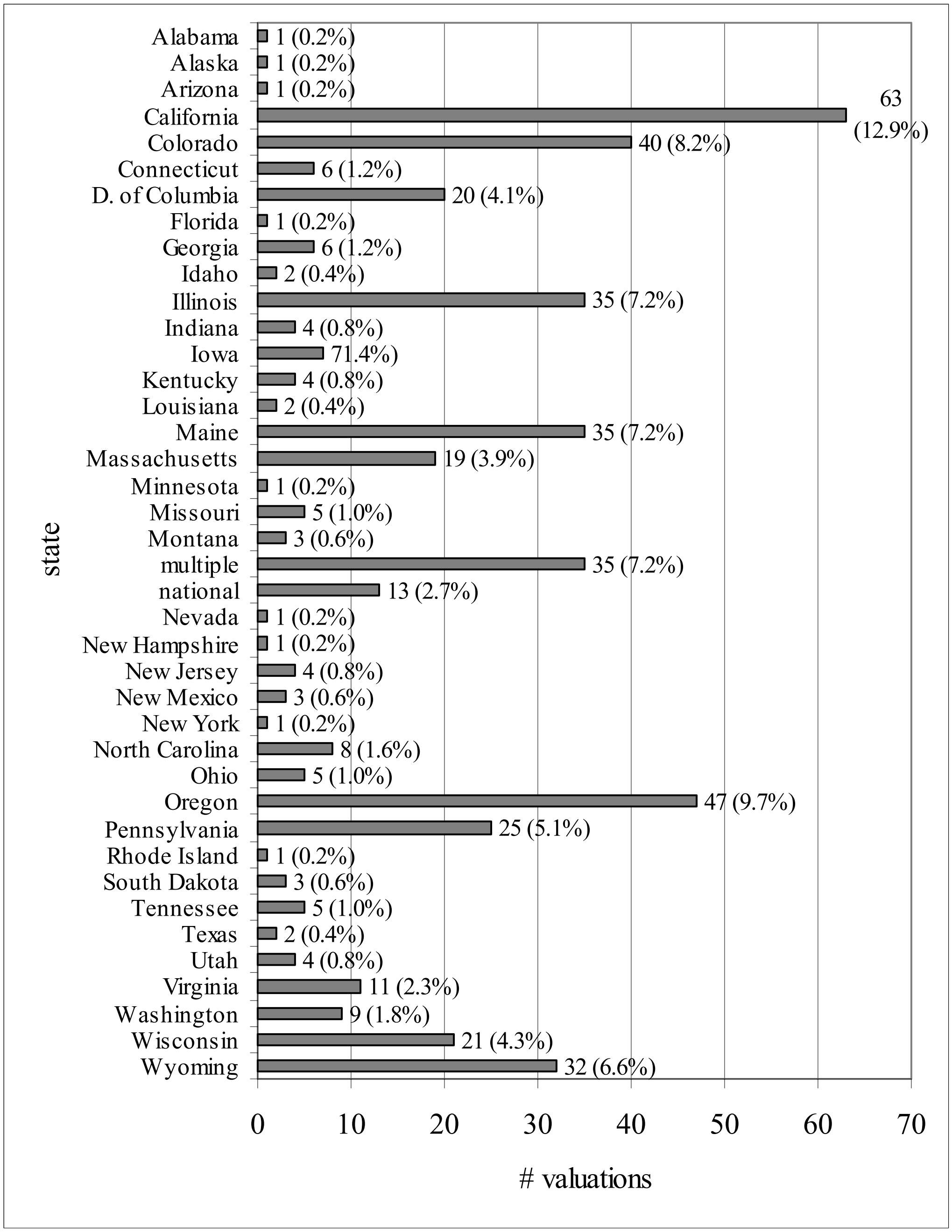


Figure 4.5: Distribution of Valuations by U.S. State

example, John Loomis is one of the most prolific scholars in this data set, appearing as author or co-author in 36 cases. His career has included significant tenures in Colorado and California, two of the better represented states in the data set.

The distribution of valuations by journal and country reveals some interesting trends, as shown in Table 4.11. For this statistical analysis, the countries were collapsed into two categories: United States and non-U.S. It shows a clear relationship with a χ^2 test statistic of 159.5, which is greater than the critical value of 13.3 for 4 d.f. and α = 0.01. Specifically, *Environmental and Resource Economics* and the *Journal of Environmental Management* disproportionately publish studies of foreign human subjects, while *Land Economics* disproportionately publish studies of U.S. human subjects. Also, although the *American Journal of Agricultural Economics* and *Water Resources Research* each published just 2 non-U.S. valuations, their expected number of non-U.S. valuations under independence is also very low, so these two journals are not as strongly associated with country as the other three are.[10]

The distribution of valuations by journal and type of environmental good is given in Table 4.12. A cursory examination of it shows one obvious relationship: *Water Resources Research* is strongly associated with valuations of water. It has no valuations of air, 4 of animals, another 4 of land, and 30 of water. This bias for water is given in the journal's title. A χ^2 test of independence of the full table results in a test statistic of 113.5, greater than the critical value of 32.0 for 16 d.f. and α = 0.01. It confirms the

[10] The 2 observations in each of these journals' non-U.S. cells violates the heuristic of a minimum of 5 per cell for the χ^2 test of independence. They were left in the analysis because their expected values were also small. However, to confirm the relationship between the other three journals and the country of study, a second χ^2 test was performed with a 2x3 contingency table leaving out AJAE and WRR. The test statistic was 134.9, far greater than the critical value of 9.2 for 2 d.f. and α = 0.01.

strong relationship between *Water Resources Research* and water. To test the relationship between the other 4 journals and the environmental goods a second χ^2 test was performed with an identical table sans *Water Resources Research*. The resulting test statistic was a much more modest 56.6, but it is still greater than the critical value of 26.2 for 12 d.f. and $\alpha = 0.01$. A closer examination of the test reveals that the *Journal of Environmental Management* is positively associated with valuations of land but negatively associated with valuations of air. Also, the *American Journal of Agricultural Economics* is positively associated with valuations of water, but negatively associated with valuations of land. Valuations published in the other two journals, *Environmental and Resource Economics* and *Land Economics*, are relatively unassociated with specific environmental goods. Furthermore, animal valuations are independent of journal.

Table 4.11: Countries Studied by Journals

	AJAE	**ERE**	**JEM**	**LE**	**WRR**	Total
U.S.	48 (8.2%)	19 (3.2%)	46 (7.9%)	338 (57.8%)	36 (6.2%)	487
Non-U.S.	2 (0.3%)	33 (5.6%)	36 (6.2%)	25 (4.3%)	2 (0.3%)	98
Total	50	52	82	363	38	585

$\chi^2 = 159.5$

Table 4.12: Environmental Goods Valued by Journals

	AJAE	ERE	JEM	LE	WRR	Total
Air	9 (1.5%)	8 (1.3%)	10 (1.6%)	121 (19.7%)	0 (0.0%)	148
Animal	9 (1.5%)	7 (1.1%)	9 (1.5%)	66 (10.7%)	4 (0.7%)	95
Land	6 (1.0%)	20 (3.3%)	41 (6.7%)	92 (15.0%)	4 (0.7%)	163
Water	25 (4.1%)	12 (2.0%)	17 (2.8%)	84 (13.7%)	30 (4.9%)	168
Other	4 (0.7%)	7 (1.1%)	5 (0.8%)	24 (3.9%)	0 (0.0%)	40
Total	53	54	82	387	38	614

5x5 χ^2 = 113.5
5x4 χ^2 (eliminating WRR) = 56.6

Next, the cross-tabulation of journals and methods of valuation is given in Table 4.13. An initial χ^2 test of independence using the full table resulted in a test statistic of 113.5, greater than the critical value of 20.1 for 8 d.f. and $\alpha = 0.01$. However, the small amount of travel cost valuations may have skewed the results. Therefore, the travel cost and hedonic pricing observations were again combined into one "weak complementarity" category and the χ^2 test was recalculated. This time, the test statistic was 55.7, still greater than the critical value of 13.3 for 4 d.f. and $\alpha = 0.01$. Thus, there is a significant relationship between journal and method of valuation. Specifically, *Land Economics* is positively associated with studies using weak complementarity methods (especially hedonic pricing), while *Environmental and Resource Economics* and the *Journal of*

Environmental Management are positively associated with studies using contingent valuation. Publications in the other two journals appear to be independent of valuation methods used.[11]

Table 4.13: Methods Published by Journals

	AJAE	ERE	JEM	LE	WRR	Total
TC	7 (1.1%)	2 (0.3%)	7 (1.1%)	3 (0.5%)	5 (0.8%)	24
HP	10 (1.6%)	10 (1.6%)	14 (2.3%)	215 (35.1%)	3 (0.5%)	252
CV	36 (5.9%)	42 (6.9%)	60 (9.8%)	169 (27.6%)	30 (4.9%)	337
Total	53	54	81	387	38	613

3x5 χ^2 = 113.5
2x5 χ^2 (combining TC and HP into WC) = 55.7

The regional distributions of valuations by environmental good and valuation method are shown in Tables 4.14 and 4.15. The χ^2 test statistic for Table 4.14 is 55.6, greater than the critical value of 13.3 for 4 d.f. and $\alpha = 0.01$. This indicates a relationship between region and good. Specifically, valuations of land is disproportionately represented outside of the U.S., while air and animal valuations are disproportionately underrepresented outside the U.S. Water values, however, receive proportionate attention

[11] These relationships apply only to studies valuing environmental goods. This set of journals also publish studies using these methods to value non-environmental goods, but such studies are not considered in this analysis. For example, the American Journal of Agricultural Economics publishes many studies using hedonic pricing to value zoning, racial make up, etc.

Table 4.14: Goods Valued by Countries

	U.S.	**Non-U.S.**	Total
Air	138 (23.6%)	10 (1.7%)	148
Animal	84 (14.4%)	6 (1.0%)	90
Land	101 (17.3%)	54 (9.2%)	155
Water	139 (23.8%)	21 (3.6%)	160
Other	25 (4.3%)	7 (1.2%)	32
Total	487	98	585

$\chi^2 = 55.6$

Table 4.15: Methods Used by Countries

	U.S.	**Non-U.S.**	Total
TC	22 (3.8%)	1 (0.2%)	23
HP	226 (38.7%)	25 (4.3%)	251
CV	239 (40.9%)	71 (12.2%)	310
Total	487	97	584

2x2 χ^2 (combining TC and HP into WC) = 18.9

inside and outside the U.S. Similarly, the χ^2 test statistic for Table 4.15 is 18.9, greater than the critical value of 6.6 for 1 d.f. and $\alpha = 0.01$. As before, travel cost and hedonic pricing methods were combined into one category of "weak complementarity" methods. The test statistic indicates a modest relationship between region and method of valuation. Specifically, outside of the U.S. contingent valuation is a disproportionately popular method. Inside the U.S. the two types of methods are proportionately represented.

Temporal Trends in Publications

The preceding analyses have focused on the descriptive statistics of the environmental valuations. Here, we examine the cases over time to find any temporal trends. For this examination, the case was chosen for analysis instead of the unit of analysis because the latter could be skewed by a few studies reporting numerous valuations of the same good. For example, one study from *Land Economics* reporting 10 different valuations of the Prince William Sound could misrepresent the amount attention *Land Economics* has spent on environmental valuation, water, and the U.S. Examining the case, however, would count that study just once for each of those categories, not 10 times. Thus, analyzing the cases speaks more closely on the practice of environmental valuation, not the valuations themselves.

First, Figure 4.6 illustrates the number of publications on four environmental goods over the last three decades. Five-year increments are used in this graph, and the next three, to smooth out the annual fluctuations in the data and better reveal the long range trends. The graph shows that land and water studies have come to dominate environmental valuation in the last decade. They have each seen a sharp rise in interest since the mid-1980s. Animal and air studies, on the other hand, have held a more steady and lower interest in the valuation literature. The relatively few articles on air value might reflect the Clean Air Act's prohibition of cost-benefit analyses in setting air quality standards (Rosenbaum, 1998, pp. 158, 163). It might also reflect the homogeneity of goods under this category. Recall that the other categories contain many different environmental goods, which may present greater unique opportunities for valuations.

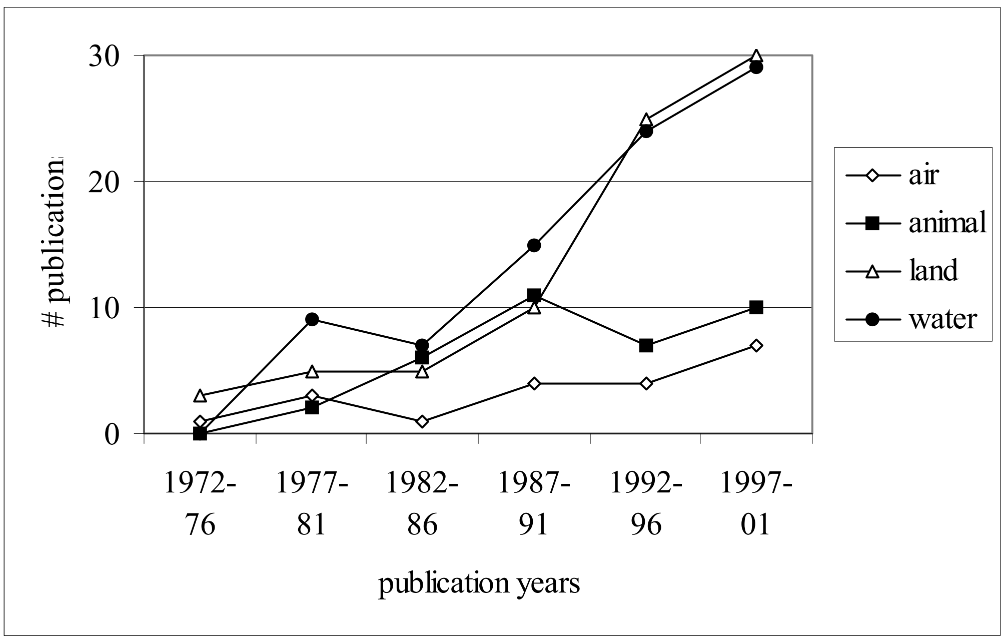


Figure 4.6: Valuations of Environmental Goods by Publication Year

Next, Figure 4.7 graphs the number of publications employing specific valuation methods against time. Again we see some clear temporal trends, and they appear to support the anecdotal evidence discussed in Chapter 2. First, we see a growth in total environmental valuation studies since the early 1970s. This coincides with the first of several Presidential Executive Orders prescribing some form of cost-benefit analysis in

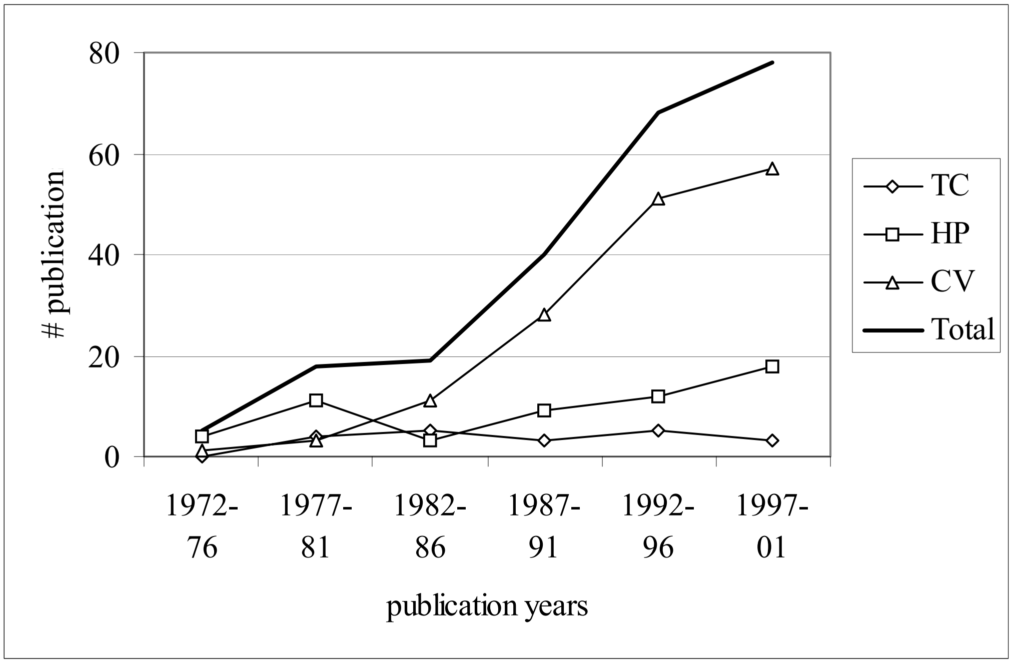


Figure 4.7: Valuation Methods Used by Publication Year

setting regulations. More specific to environmental issues, the beginning of this trend also coincides with the Water Resources Council's prescriptions for environmental valuation in 1973. The sharp rise in studies in the late 1980s might be due in part to President Reagan's Executive Order 12291 in 1981 recommending only those regulations whose benefits can be shown to exceed their costs. Second, we see that the growth in environmental valuation studies since the mid-1980s was due primarily to a growth in the use of contingent valuation. Between 1987 and 1991, contingent valuation came to represent a majority of the valuation studies in these five journals. This period coincides with both the publication of Mitchell and Carson's (1989) seminal text on the method, and the *Exxon Valdez* oil spill which brought much attention to the method. Four years later the National Oceanic and Atmospheric Administration published the findings of its blue ribbon panel on contingent valuation (Arrow et al., 1993), which gave a qualified endorsement of the method and further popularized it. During this same period, the number of environmental valuation studies employing the hedonic pricing rose just slightly. The number employing the travel cost method remained steady but few. Indeed, while the travel cost method can be used to value changes in environmental goods, its niche is in valuing outdoor recreation and its popularity lies in the latter, not the former.

Figure 4.8 charts the numbers of environmental valuation studies per journal against time. Each of the five journals has seen a long-term growth in the number of valuation studies it publishes. *Land Economics* consistently publishes the most, although the two new journals – *Environmental and Resource Economics* and *Journal of Environmental Management* – have quickly become rivals for the title.

Finally, Figure 4.9 illustrates the numbers of studies done in the United States and elsewhere against time. It shows a 5 to 10 year gap in popularity between the U.S. and all other countries, with the U.S. leading the efforts in environmental valuation. However, as shown in Figure 4.4, no single country has even approached the popularity of environmental valuation that is seen in the U.S.

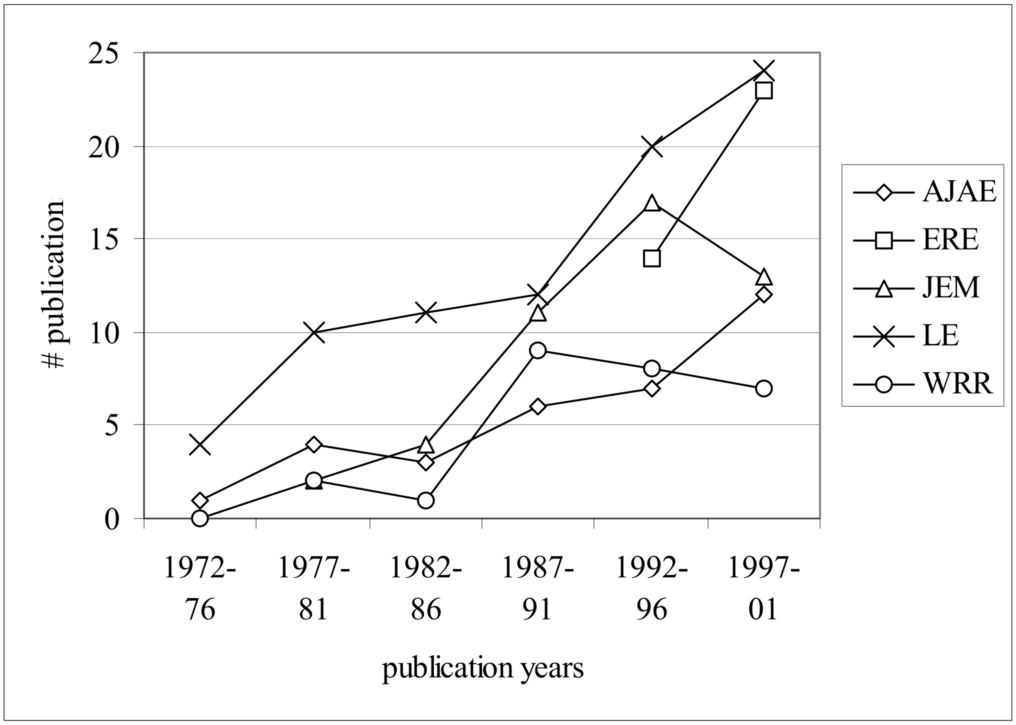


Figure 4.8: Journal Publications by Publication Year

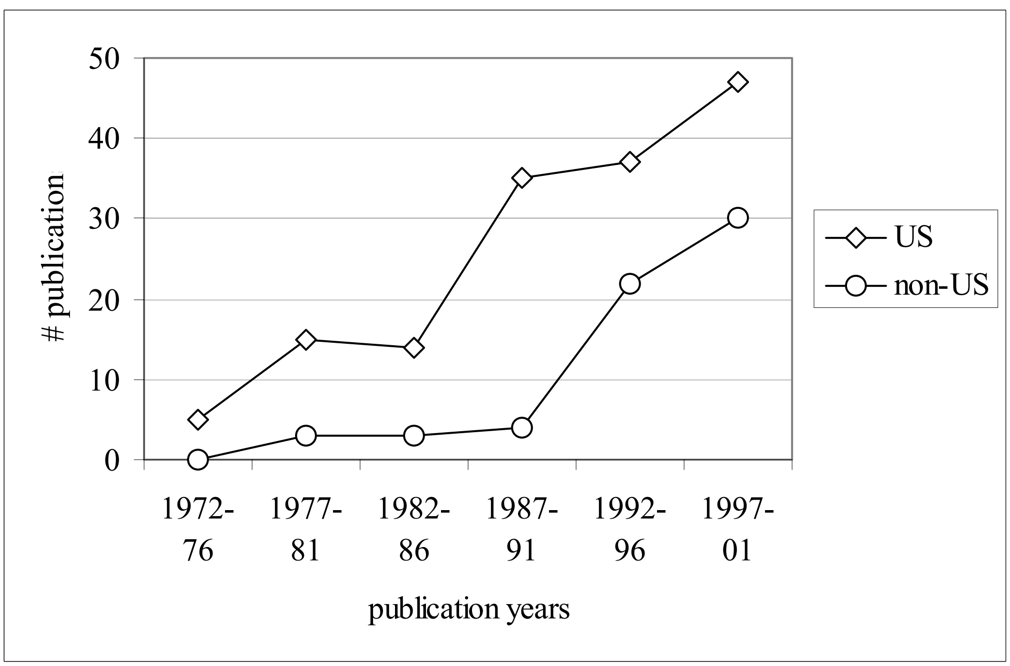


Figure 4.9: Country of Study by Publication Year

Summary

To summarize the findings described in this chapter, the 228 cases and 614 units of analysis in this data set of environmental valuation studies exhibited the following major characteristics:

- Water, land, and air are well represented in the valuation studies. Animals are too, but the diversity of species valued make their comparability, and the transferability of their results, questionable. Other environmental goods such as plants, toxics, and wastes, have received relatively little attention in environmental valuation.

- Contingent valuation is the most popularly used valuation method, representing a majority of the valuations in this data set. Although it is the youngest of the three methods, it is also the most versatile, which contributes to its popularity. Furthermore, there is evidence that the federal government endorsement of it has increased its used.

- Different methods tend to be used in different situations. Specifically, contingent valuation tends to be used with qualitative descriptions of holistic changes in environmental goods. While it is used to value all types of goods, it is especially favored when valuing animals. Hedonic pricing tends to be used with quantitative descriptions of incremental changes. Its use has been limited to valuing water, land, and especially air. The travel cost method is rarely used to value changes in environmental goods.

- The vast majority of valuation studies have involved human subjects in the United States. This may be a reflection of both the location of the scholars publishing valuation studies, as well as government promotion of cost-benefit analysis and environmental valuation there.

- There is evidence that certain journals disproportionately publish valuations from different places, of different goods, or using different methods, as shown in Table 4:16.

Table 4.16: Journals' Associations with Types of Valuations

	AJAE	ERE	JEM	LE	WRR
Country					
U.S.		–	–	+	
non-U.S.		+	+	–	
Good					
Air			–		–
Animal					
Land	–		+		
Water	+				+
Other					
Method					
WC		–	–	+	
CV		+	+	–	

+ journal is positively associated with valuations with this characteristic
– journal is negatively associated with valuations with this characteristic

- Although environmental valuation outside of the U.S. has received little attention, the attention it has received is heavily focused on land values, and not so on air or animal values. Also, valuations outside the U.S. are more likely to employ contingent valuation than hedonic pricing or the travel cost method.

- There has been a sharp rise in environmental valuation studies since the mid-1980s. The rise has been characterized by studies on land and water goods, and studies employing contingent valuation. The rise coincides with a few events – especially the maturation of contingent valuation – that have helped make environmental valuation popular.

CHAPTER 5

RESULTS

This chapter reports the analysis of the valuations of environmental changes. As discussed in Chapter 3, the individual valuations were converted to standardized effect sizes to provide a common measure of the dependent variable and to facilitate comparability of results. Thus, this chapter begins with an overall descriptive assessment of the effect sizes and the indications of the presence of moderating variables. It then presents the tests for the individual moderating variables hypothesized in Chapter 3, including type of good, magnitude of change, description of change, method of valuation, and secondary variables of interest. These tests include reports of the relative effects of all the confirmed moderating variables in meta-regression models of the data set. Next, the unexplained, remaining variance in the dependent variable and the limits of the results are discussed.

Effect Sizes

For each observation with sufficient reported information an effect size was calculated by the general equation $d = (Y_E - Y_C)/S$ (see Chapter 3). A few studies valued a degradation of an environmental good, not an improvement, such as when respondents were asked for their willingness to accept a reduction in air quality, or when a hedonic pricing model included a measure of pollution as an independent variable. In these cases, the changes in the environmental goods were transformed into improvements. That is,

Y_E was recorded as the value of an improved environmental state compared to that for Y_C, so ($Y_E - Y_C$) was always the change in value for an improved condition.

The effect sizes were corrected for sampling error and combined, following Hunter and Schmidt's (1990) procedures[1]. The average effect size weighted by sample size, known as the mean effect size or the observed effect size, is

$$\mathrm{Ave}(d) = D = [\Sigma(n_i d_i)]/[\Sigma n_i] = 6.90,$$

where n_i and d_i are the sample size and effect size for observation i, respectively. This number indicates that, based on this data set, the average change in value of an environmental good due to its improvement is 6.90 standard deviations. Notice that the effect size is does not specify an absolute change in value, but a relative change in units of standard deviations. The advantages of this approach, as discussed in Chapter 2, are that it allows the inclusion of more studies using different outcome measures, and it accounts for the variability within the results of each study. These two factors allowed me to make broader conclusions about the body of environmental valuation studies.

Of course, the summary value overlooks the variance *among* the studies generating this data set, such as differences in treatment (improvement in the environmental good), subject (the environmental good), etc. These factors are addressed in the rest of this chapter. Still, the direction and order of magnitude of the mean effect

[1] The pure lack of reported study artifacts limited the corrections to what Hunter and Schmidt (1990) called the "bare bones" (p. 100): correcting only for sampling error variance. Other study artifacts – such as measurement errors in the independent and dependent variables, study attrition, reporting errors, etc. – were rarely, if ever, reported. In fact, many studies examined for this research reported insufficient information to even calculate effect sizes or sampling error variance (e.g., they did not report standard deviations or sample sizes). Other meta-analysts of environmental valuation have encountered a similar lack of reported artifacts (Carson, Flores, Martin, & Wright, 1996).

size is instructive. It tells us that, assuming the validity of the valuation methods (an assumption questioned in this research), people do place a significant monetary value on improved conditions of the environment. This certainly concurs with the claim that a broad ideology of "environmentalism" – one at least as encompassing as the environmental goods represented in this data set – has become a "core" value in the United States (Kempton, Boster, & Hartley, 1999, p. 214), the location of the bulk of the cases in this data set.

Continuing with the sampling error analysis, the observed variance of the effect size weighted by sample size is

$$\mathrm{Var}(d) = \{\Sigma[n_i\,(d_i - D)^2]\}/[\Sigma n_i] = 52.68.$$

Finally, the sampling error variance – the amount of variance in the effect size that can be attributed to sampling error – is

$$\mathrm{Var}(e) = [(N-1)/(N-3)][4/N][1+(D^2/8)] = 0.0136$$

where N, the average sample size, is the total sample size, $T = \Sigma n_i$, divided by the total number of effect sizes examined, K.

$$N = T/K = 776295/434 = 1788.70.$$

This result indicates that sampling error accounts for very little of the variance in effect size. Specifically, when the data set is considered as a whole, sampling error accounts for just 0.03% of the total variance.[2]

[2] Var(e)/Var(d) = 0.0136/52.68 = 0.03%.

The relative magnitude of Var(e) also gauges the homogeneity of the total sample. Hunter and Schmidt's (1990) "75% rule" is a heuristic which states that if more than 75% of the variance in effect size is shown to be due to study artifacts (such as sampling error), then the remaining 25% is likely to also be due to study artifacts, and one can conclude that the population is homogeneous. If, however, less than 75% of the variance is due to study artifacts, then the population studied may be heterogeneous and there may be moderators explaining the residual variance. As stated above, sampling error accounts for just 0.03% of the total variance in effect sizes, far less than 75%. Of course, only sampling error was accounted in this calculation, and much more the variance might actually be explained by other study artifacts that were unaccountable in this meta-analysis. The χ^2 (chi-square) test of heterogeneity[3] was also performed to confirm heterogeneity (Hunter & Schmidt, 1990). The test statistic Q is

$$Q = K\ \mathrm{Var}(d)/\mathrm{Var}(e) = 434(52.68/0.0136) = 1681221,$$

which is far greater than the critical value at $\alpha = 0.01$. Thus, heterogeneity is confirmed and moderating variables may be present.

---

[3] The χ^2 test of heterogeneity should not be confused with the χ^2 test of independence used in Chapter 4. Both are based upon a χ^2 distribution, hence the similar names. However, the χ^2 test of heterogeneity measures the heterogeneity, and homogeneity, of the samples combined in a meta-analysis. Formally, the null and alternative hypotheses are

H_0: the population is homogeneous

H_1: the population is not homogeneous.

The null hypothesis is rejected if the test statistic is greater than the critical value (defined by the degrees of freedom in the sample, and a selected level of significance, α), otherwise the null cannot be rejected. A rejection of homogeneity is an indication that there may be moderating variables explaining the high variance.

Hypothesized Moderators

The four primary hypothesized moderators were tested first: type of good, magnitude of change, description of change, and method of valuation. The operational measures of each of these variables created categorical data, so the meta-analysis of each consisted of dividing the data set by the discrete values of the hypothesized moderator, calculating the mean effect size of each subset, and comparing these means. Significant differences in the means of the subsets demonstrate the moderating effect of the variable in question.

Environmental Good

Thus, to test the moderating effect of the type of environmental good, the data set was divided into the six major categories described in Chapter 4: air, animal, land, plant, toxic, and water. Table 5.1 summarizes the effect sizes for the goods and their corresponding analyses. While the average effect size of all six goods combined is 6.93, the average effect sizes of the individual goods range from 1.37 standard deviations for animal to 18.90 for air. These statistics indicate that an improvement in the condition of animals results in a 1.37 standard deviation increase in their value, and an improvement in the condition of air results in a 18.90 standard deviation increase in value. The effect sizes do not convey the absolute magnitudes of the values of environmental goods, only their changes in value relative to their own baseline values. Thus, a change in the condition of air results in a greater change in its value than that for a change in the conditions of animals.

Table 5.1: Mean Effect Sizes of Environmental Goods

	sample size T	**# of effect sizes K**	**mean effect size D**	**Var(d)**	**Var(e)**	**Var(e) / Var(d)**	**χ^2 test statistic Q**	**99% C.I. for δ**
Combined*	771525	423	6.93	52.97	0.01	0.03%	1666988	6.00 7.85
Air	217189	135	18.90	78.33	0.11	0.14%	94858	16.10 21.71
Animal	16363	68	1.37	3.31	0.01	0.18%	37351	0.78 1.96
Land	442980	112	1.38	16.24	0.00	0.00%	4964084	0.39 2.37
Plant	4962	2	1.54	0.02	0.00	3.22%	62	-5.02 8.10
Toxic	2419	4	7.49	118.65	0.05	0.04%	10022	-24.98 39.96
Water	87612	102	6.63	186.27	0.03	0.01%	724916	3.09 10.16

*Includes only the 6 categories of goods in the table, not the entire data set.

A one-way analysis of variance (ANOVA) of the effect sizes resulted in F = 37.29[4], demonstrating a significant difference in the average effect sizes of the goods.[5] To determine *which* of the goods are significantly different from others, Fisher's least significant difference (LSD) test was performed. The LSD test is a modification of the t-

[4] The one-way ANOVA tests the differences between the mean values of more than 2 groups. Formally, the null and alternative hypotheses are

H_0: $b_1 = b_2 = \ldots = b_k$ (the means of all the groups are equal)

H_1: not all b_k are equal.

The null hypothesis is rejected if the test statistic, F, is greater than the critical value (defined by two measures of the degrees of freedom in the sample, and a selected level of significance, α), otherwise the null cannot be rejected. A rejection of the null hypothesis indicates significant differences between at least two of the means, and thus the moderating effect of the grouping variable, which in this case is environmental good.

[5] In this case, the critical value of F for 5 degrees of freedom in the numerator and 417 degrees of freedom in the denominator, and at a level of significance of $\alpha = 0.01$, is 2.24. The test statistic is 37.29. Thus, with 99% confidence the null hypothesis is rejected.

test for the difference in means between two independent samples.[6] The modification accounts for the multiple pairs in an ANOVA and minimizes the chances of a type 1 error.[7] Table 5.2 summarizes the results of the LSD test. It reveals that the mean effect sizes for air and water are significantly different from all the other types of goods except plant and toxic. Plant and toxic, however, had such few observations in the data set (2 and 4, respectively) that their differences with the others could not be statistically confirmed. Still, the test supports what appears to be evident in Table 5.1, that air and water have significantly larger average effect sizes than animal and land, and air has a significantly larger average effect size than water.

Table 5.2: t-Statistics from Fisher's LSD Test of the Differences Between the Mean Effect Sizes of Pairs of Environmental Goods

	Air	**Animal**	**Land**	**Plant**	**Toxic**	**Water**
Air	---					
Animal	10.19*	---				
Land	11.85*	0.00	---			
Plant	2.11	-0.02	-0.02	---		
Toxic	1.94	-1.03	-1.04	-0.59	---	
Water	8.08*	-2.92*	-3.31*	-0.62	0.15	---

* Significantly different at $\alpha = 0.01$

[6] In a simple t-test of the difference in the means of two independent samples, the means from two groups are compared. Formally, the null and alternative hypotheses are

H_0: $\mu_1 = \mu_2$ (the means are equal)

H_1: $\mu_1 \neq \mu_2$ (the means are not equal).

The null hypothesis is rejected if the test statistic, t, is outside the two critical values defined by the degrees of freedom in the samples and the level of significance, α. Fisher's LSD test is a modified series of t-tests for every pair-wise combination of groups in a one-way ANOVA. In this case, with six groups, fifteen different pairs were tested by Fisher's procedures.

[7] A type 1 error is committed when a null hypothesis is rejected when it is actually true. The acceptable level of a type 1 error is the level of significance, α, which is set by the researcher.

While it should be no surprise that changes in different environmental goods are valued differently, there has been little said of their relative values. One study in the data set concluded that valuations of land changes are highly variant (Willis & Garrod, 1993), which would result in lower effect sizes[8]. The results in Table 5.1 confirm this, and the variety of land types and features valued might explain the high variance. Table 4.1 lists 21 different land types, ranging from Superfund sites to wilderness, valued in this data set. With so many different land types valued, the high variance in valuations is reasonable. The same can be said for the animal category. At least 26 species are included in the data set, ranging from black flies to whales, giving the data high variance and a low average effect size. On the other hand, the air category is homogenous, with all valuations in the data set aimed at air quality. Certainly there were different measures of air quality used in studies, but the object of valuation was consistently air quality. It is thus reasonable to expect less variance, and a larger average effect size, for this homogenous category. Indeed, it has the largest average effect size of all the goods. The water category is in the middle, with an average effect size between those of air and land or animal. While it contains 25 subcategories, ranging from stream to ocean, people's valuations of them are not as variant as those for land or animal categories.

Table 5.1 also shows two confidence intervals that include zero. The confidence intervals for the average effect sizes for plant and toxic categories have negative lower bounds and positive upper bounds. This is due to the low number of observations for

[8] Higher variance *within* each unit of analysis lowers its resulting effect size, because variance is captured in the denominator of $d = (Y_E – Y_C)/S$. The variances reported in Tables 5.1, 5.3 – 5.5, and 5.7 – 5.10 are those *among* the units of analysis.

each of those goods. For all of the other categories of goods, the confidence intervals are completely positive, indicating positive valuation of environmental improvements.

Finally, the results in Table 5.1 also suggest the presence of additional moderating variables, even after accounting for the type of environmental good. This is demonstrated in the Var(e)/Var(d) percentages, all far less than Hunter and Schmidt's 75% rule. They are also confirmed by the values of the Q statistics for all of the goods. Thus, while the type of good is confirmed to be a moderating variable, others are likely to exist as well.

Magnitude of Change

As discussed in Chapters 3 and 4, the magnitudes of changes in the environmental goods were coded as incremental and holistic. The data set was split into these two categories and the average effect size was calculated for each. The results are shown in Table 5.3, and they are quite clear. The average effect size for a holistic change in an environmental good is 0.84 standard deviations, while that for an incremental change is 9.46. A t-test[9] for difference between these two values resulted in test statistic of t = 10.57. With 99% confidence there is a significant difference between the effect sizes of holistic and incremental environmental changes.[10]

---

[9] The t-test for the difference between two means from independent samples assumes equal variances of the two samples. Because this cannot be assumed in this case (the variance of one is two orders of magnitude larger than the other), separate variances were used in the calculation of the test statistic, and the smaller of n_1-1 and n_2-1 was used as the degrees of freedom. This adjustment is a conservative estimate that reduces the chances of a type 1 error.
[10] At $\alpha = 0.01$ and 430 degrees of freedom, the critical value of t is 2.6 for a two-tail test.

Table 5.3: Mean Effect Sizes of Magnitudes of Changes

	sample size T	# of effect sizes K	mean effect size D	Var(d)	Var(e)	Var(e) / Var(d)	χ^2 test statistic Q	99% C.I. for δ
Combined	750481	432	7.05	99.54	0.02	0.02%	2583889	5.81 8.30
Holistic	209628	216	0.84*	8.84	0.00	0.05%	424623	0.32 1.37
Incremental	540853	216	9.46*	134.69	0.02	0.01%	1493123	7.40 11.52

* Significantly different at $\alpha = 0.01$

At first glance, however, the difference is somewhat counterintuitive. Incremental changes in the environment resulted in larger relative changes in valuations than holistic changes did. This may be because incremental changes are associated with quantitative measures (see Table 4.4), and quantitative measures have a better common understanding (i.e., a more reliable measure) among subjects than qualitative measures. Similarly, the smaller average effect size for holistic changes may reflect a larger variance of responses. For example, responses to changes in air quality from "unhealthy" to "healthy" may be more variant than responses to changes measured in "miles of visibility."

The Var(e)/Var(d) percentages and the Q test statistics in Table 5.3 strongly suggest the presence of other moderating variables, consistent with the findings from the analysis of the types of goods as a moderating variable. It was expected that the analyses of the rest of the potential moderating variables would also suggest the presence of other moderating variables, and indeed they did.

Description of Change

The changes in the environmental goods were also coded by their quantitative or qualitative descriptions, and these types of descriptions were tested for a moderating effect on valuations. As discussed in Chapter 3, there are logical reasons to expect a difference in valuations between the two types of descriptions, but it is not clear which might be larger. Table 5.4 summarizes the meta-analysis of the data split by quantitative and qualitative measures of change, and clearly answers this question. The average effect size for qualitative descriptions of changes is 0.96 standard deviations, while that for quantitative descriptions is 8.69. The difference is statistically significant at a 99% confidence level.[11] Thus, this data demonstrate that quantitative descriptions of environmental change produce significantly larger effect sizes than qualitative descriptions. This conclusion supports the hypothesis that quantitative measures produce less variant measures of value than qualitative measures, resulting in larger effect sizes. It also undermines the hypothesis that qualitative descriptions of changes in environmental goods are better understood (i.e., more reliable) than quantitative descriptions.

These results support recent evidence from studies focusing on measures of water quality. Michael, Boyle, and Bouchard (2000) tested nine measures of water clarity in a hedonic pricing model of lakefront property in Maine. They found significant differences in the values from the different measures of clarity and concluded that the differences were large enough to affect water policy decisions. Poor, Boyle, Taylor, and Bouchard

---

[11] The difference in means test for independent samples resulted in $t = 10.61$. Using the smaller of n_1-1 and n_2-1 as the degrees of freedom, the critical value of t at $\alpha = 0.01$ is 2.6. The test statistic is larger, so the null hypothesis of no difference in means is rejected.

(2001), in another hedonic pricing study of the value of water clarity showed that objective, quantitative measures of water clarity were more reliable than subjective, qualitative ones.

Table 5.4: Mean Effect Sizes of Descriptions of Changes

	sample size T	# of effect sizes K	mean effect size D	Var(d)	Var(e)	Var(e) / Var(d)	χ^2 test statistic Q	99% C.I. for δ
Combined	776295	434	6.90	100.74	0.02	0.02%	2809997	5.65 8.15
Qualitative	179917	162	0.96*	9.63	0.00	0.04%	387353	0.34 1.59
Quantitative	596378	272	8.69*	128.22	0.02	0.02%	1829597	6.90 10.48

* Significantly different at $\alpha = 0.01$

However, the significant difference might be explained by a spurious effect from the method of valuation. As Table 4.8 shows, qualitative descriptions of environmental changes are associated with contingent valuation, and quantitative descriptions are associated with hedonic pricing. If the method of valuation is shown to have a significant moderating effect on valuation too, then the relative effects of type of description of change and method of valuation can be tested together in meta-regression analyses. These analyses are next.

Valuation Method

The primary variable of interest in this study is the method of valuation. The data set was split into three groups, one for each of the methods studied, and the average effect size of each of the subsets was calculated and compared with the others. Table 5.5 summarizes the results. The average effect sizes ranged from 3.16 standard deviations for contingent valuation to 7.23 standard deviations for hedonic pricing. A one-way ANOVA, testing the difference in the average effect sizes of the 3 methods, resulted in a test statistic of $F = 31.01$. Thus, with 99% confidence the null hypothesis of the equality of the means is rejected.[12] To identify which average effect sizes are significantly different from the others, Fisher's least significant difference test was again employed. As Table 5.6 shows, the average effect size from contingent valuation is significantly different from those of the hedonic pricing and travel cost methods. The difference is more significant between contingent valuation and hedonic pricing than between contingent valuation and the travel cost method. However, this is due in part to the small number of observations employing the travel cost method. No significant difference was found between the average effect sizes of travel cost and hedonic pricing studies.

Comparing contingent valuation against hedonic pricing, the results indicate that on average hedonic pricing of changes in environmental goods results in measures of value that are over twice those of contingent valuation studies. This is quite similar to results from an earlier meta-analysis of 83 studies that contained both contingent valuation and travel cost or hedonic pricing estimates for the same goods (Carson, Flores, Martin, & Wright, 1996). It concluded that contingent valuation estimates run about 20% to 30% lower than travel cost estimates and 40% less than hedonic pricing estimates. In

[12] The critical value of F for 2 degrees of freedom in the numerator and 430 degrees of freedom in the denominator, at $\alpha = 0.01$, is 4.7. The test statistic is larger than the critical value. Thus, the null hypothesis is rejected.

Table 5.5: Mean Effect Sizes of Methods of Valuation

	sample size T	# of effect sizes K	mean effect size D	Var(d)	Var(e)	Var(e) / Var(d)	χ^2 test statistic Q	99% C.I. for δ
Combined	775960	433	6.90	110.19	0.02	0.01%	3070230	5.60 8.21
TC	8197	16	7.00	67.50	0.06	0.08%	19362	0.75 13.24
HP	705963	204	7.23	116.56	0.01	0.01%	2730001	5.28 9.18
CV	61800	213	3.16	43.06	0.03	0.07%	293925	2.00 4.32

Table 5.6: t-Statistics from Fisher's LSD Test of the Differences Between the Mean Effect Sizes of Pairs of Valuation Methods

	TC	HP	CV
TC	---		
HP	0.13	---	
CV	2.13**	5.97*	---

* Significantly different at $\alpha = 0.01$

** Significantly different at $\alpha = 0.05$

the current data set, which includes much more varied studies, contingent valuation estimates run about 55% lower than travel cost and hedonic pricing estimates.

However, proponents of these valuation methods could dispute these conclusions because the temporal scope of the data set includes early studies published before the development of the contemporary practices for two of the methods. As discussed in Chapter 1, the hedonic pricing method underwent significant refinements through the 1970s, and contingent valuation underwent significant refinements through the 1980s. The data set used for this research includes studies published since 1970. One could reasonably argue, then, that the differences in the average effect sizes between the methods of valuation may be affected by the inclusion of older studies that are less reliable than more contemporary studies.

This argument was tested through two additional meta-analyses of the data set in which the year of publication was the hypothesized moderating variable. This variable was tested on the set of hedonic pricing studies and separately on the set of contingent valuation studies. The set of travel cost studies was not tested because the temporal frame of the data set is completely after the major developments in the method.

First, the hedonic pricing studies were analyzed. A total of 204 effect sizes were calculated from the studies employing hedonic pricing. These were divided into two groups: those published in 1980 or earlier, and those published after 1980. If the studies published after 1980 produce a significantly greater average effect size (i.e., less variant) than those published up to 1980, then it could be argued that the more recent studies are more reliable than the older ones. Table 5.7 shows the results of the analysis. There were only 23 observations in the older group, with an average effect size of 3.40 standard deviations. There were 181 observations in the younger group, with an average effect size of 7.26. A t-test of the difference between the two means resulted in a test statistic of $t = 1.63$. At a level of significance of $\alpha = 0.01$, the difference between the means is insignificant.[13] This is also evident in the overlapping confidence intervals for the average effect sizes. In other words, there is insufficient evidence to demonstrate a statistically significant difference in the average effect sizes of hedonic pricing studies done before and after 1980. In this data set, the method has been reliable over the years.

The same cannot be said about the contingent valuation method. The 213 effect sizes calculated from contingent valuations were divided into two groups: the 101 published in 1993 or earlier, and the 112 published after 1993. That was the year the National Oceanic and Atmospheric Administration's report on contingent valuation (Arrow et al., 1993) was published, and its publication is regarded as the event that legitimized the method in the federal government and standardized its practices (Carson, 1997). As Table 5.8 shows, the average effect size of the older studies is 1.24 standard deviations, while that of the younger studies is 4.33. A t-test of the difference between these two means resulted in a test statistic of $t = 4.00$, confirming a significant difference.[14] On average, effect sizes calculated from contingent valuations published in 1993 or earlier are 71% less than those calculated from studies published after 1993. This confirms that part of the difference between the average effect sizes of contingent

[13] Because the variances of the two subgroups are nearly equal, the degrees of freedom was calculated as d.f. = $(n_1 + n_2 - 2) = 202$. At $\alpha = 0.01$, the critical value of t is 2.6. (At $\alpha = 0.10$ it is 1.7.) The test statistic is less than this, so the null hypothesis of the equality of the means cannot be rejected.

[14] Separate variances were used in the calculation of the test statistic. The degrees of freedom were calculated as the smaller of $n_1 - 1$ and $n_2 - 1$. At $\alpha = 0.01$, the critical value of t is 2.6. Thus, with 99% confidence, the null hypothesis of the equality of the two means is rejected.

Table 5.7: Mean Effect Sizes of Hedonic Pricing Studies by Year of Publication

	sample size T	# of effect sizes K	mean effect size D	Var(d)	Var(e)	Var(e) / Var(d)	χ^2 test statistic Q	99% C.I. for δ
Combined	705963	204	7.23	116.44	0.01	0.01%	2727112	5.28 9.18
HP ≤ 1980	5883	23	3.40	114.04	0.04	0.03%	68127	-2.89 9.69
HP > 1980	700080	181	7.26	116.46	0.01	0.01%	2684237	5.19 9.33

Table 5.8: Mean Effect Sizes of Contingent Valuation Studies by Year of Publication

	sample size T	# of effect sizes K	mean effect size D	Var(d)	Var(e)	Var(e) / Var(d)	χ^2 test statistic Q	99% C.I. for δ
Combined	61800	213	3.16	40.80	0.03	0.08%	278522	2.03 4.29
CV ≤ 1993	23422	101	1.24	2.99	0.02	0.69%	14590	0.79 1.69
CV > 1993	38378	112	4.33	63.88	0.04	0.06%	182050	2.36 6.31

valuations and hedonic pricing studies is due to the evolving, and apparently improving, reliability of contingent valuation.

Still, when the hedonic pricing results are compared with only the post-1993 contingent valuation results, a significant difference remains. The t-test for the difference between these average effect sizes (7.23 for hedonic pricing and 4.33 for post-1993 contingent valuation) resulted in a test statistic of $t = 2.71$.[15] The difference between the means is statistically significant. After 1993, effect sizes from contingent valuations were on average 40% less than those from hedonic pricing. This result matches that of the earlier, more restricted, meta-analysis (Carson, Flores, Martin, & Wright, 1996). Thus, while the evidence suggests that the reliability of contingent valuation has improved over the years, there remains a divergence between its results and those of the travel cost and hedonic pricing methods.

What might contribute to this residual disparity? One clear factor is the variety of implementation methods of contingent valuation that have been shown to affect results. At least four design characteristics are relevant. First is the method of eliciting responses. Dichotomous choice questions (with and without follow-up questions), discrete choice payment cards, and open-ended questions have been shown to produce significantly different results (Boyle, MacDonald, Cheng, & McCollum, 1998; Brown, Champ, Bishop, & McCollum, 1996; Kristrom, 1997; Ready, Buzby, & Hu, 1996; Ready, Navrud, & Dubourg, 2001; Scarpa & Bateman, 2000). Second, different payment vehicles have been shown to affect results (Morrison, Blamey, & Bennett, 2000). Lump

[15] Again equal variances was not assumed. At $\alpha = 0.01$ and 111 degrees of freedom the critical value is $t = 2.6$. Thus, with 99% confidence, the null hypothesis of the equality of means is rejected.

sums tend to underestimate value while payment schedules tend to grossly overestimate value (Stevens, DeCoteau, & Willis, 1997), and the differences between willingness to pay and willingness to accept have both theoretical explanations (Mansfield, 1999) and empirical evidence (Adamowicz, Bhardwaj, & Macnab, 1993; Knetsch, 1990; Kolstad & Guzman, 1999). Third, survey research has shown that the mode of elicitation affects results of contingent valuations. The results coming from phone, mail, and in-person surveys have been shown to be significantly different (Ethier, Poe, Schulze, & Clark, 2000; Mannesto & Loomis, 1991). Even the number of reminders needed to procure responses to mailed contingent valuation surveys has been shown to affect responses (Dalecki, Whitehead, & Blomquist, 1993). Fourth, the quality and quantity of information provided on the environmental goods and hypothetical markets have been shown to affect results (Boyle, Bergstrom, & Poe, 2001; Cummings & Taylor, 1998; Hanley, Spash, & Walker, 1995; Randall, Hoehn, & Brookshire, 1983; Whitehead & Blomquist, 1991).

This last factor, information bias, is dramatically exemplified with a few cases from the data set. Consider, for example, the valuation of possums. How should humans impart value on these animals, which, while fulfilling their roles in their ecosystems, are considered ugly pests by many humans? In one contingent valuation study, respondents were told of possums' ecological role and their threatened status in a local environment. They were then asked for their willingness to accept compensation for the complete loss of the species in the local environment (Jakobsson & Dragun, 2001). The question drew strong protest responses because respondents objected to the idea of being compensated for the deliberate loss of a wild species. Instead, respondents expressed a willingness to pay to preserve the species. In another contingent valuation study, however, respondents were asked to value "possum control" (Kerr & Cullen, 1995). The quality of a local forest was described as a function of the possum population. Reductions in the possum population would increase the quality of the forest. Respondents in the study expressed a willingness to pay for a reduction in the possum population. Other environmental goods have similarly been valued positively and negatively, based upon their contextual presentation. Plant quantity, for example, can be positive when presented as the number of trees in a park (Goldar & Misra, 2001), and negative when presented as lake coverage by an aquatic plant (Messonnier, Bergstrom, Cornwell, Teasley, & Cordell, 2000). These wide ranges in values for similar environmental goods increase the variance among their effect sizes, and thus decrease their average effect size. Negative valuations of environmental goods are not unique to contingent valuation studies. Several hedonic pricing studies in the data set also produced negative effect sizes for environmental improvements. However, contingent valuation is unique in that negative and positive valuations can be directly affected by the researchers through the framing of the environmental good. This vulnerability to information bias creates greater variability, and less reliability, among their results.

Another form of information bias relates to the described scope of the good. While environmental valuation techniques have been developed to measure values of environmental externalities of public policies, environmental goods have externalities of their own that are often overlooked when their values are measured in economic terms. For example, improvements in water qualities or quantities can affect habitats and land as well as human consumers of water, but oftentimes these externalities are not described in the hypothetical markets. Another example is air quality. Its improvement can also affect water and land qualities. Acid rain is a good example of this. Reductions in sulfur dioxide in the air is positively valued by humans for its direct health benefits, but it also has indirect benefits in water and land quality (acidity of rain and runoff), ecosystems, etc. Studies valuing air quality, however, may neglect these externalities. Without explicitly including or excluding such externalities in a contingent valuation, the scope of the valuation is unclear and the variation in the results can reasonably be expected to be large. In hedonic pricing and travel cost methods, the scope of the environmental good is inferred by the researcher but not described to the subjects, so information biases do not exist.

With all these design variables shown to affect results of contingent valuations – especially information bias - the variance among the results can be expected to be greater, and the average effect size can be expected to be smaller.

The lower average effect size of the contingent valuation studies might also be explained by the method's susceptibility to cognitive problems and strategic replies of respondents, as defined in Chapter 1. Cognitive problems include respondents' abilities to understand the hypothetical market and quantify their values for the goods. An experiment that compared the contingent valuation of a real market good with actual market behavior found discrepancies in value that ranged from 1% to 124% (Kealy, Dovidio, & Rockel, 1988). It concluded that even for well defined, familiar market goods, contingent valuation results diverged from actual market behavior. Even greater divergences can be expected for less well defined, less familiar, nonmarket goods such as environmental goods (Champ & Bishop, 2001). Strategic replies are deliberate overstatements and understatements of value intended to influence or protest outcomes. Furthermore, such strategic responses are not limited to obvious outliers, such as zero bids that are often removed from analyses as protests. Positive bids included in analyses may also hold protests of some kind (Jorgensen, Syme, Bishop, & Nancarrow, 1999).

The disparities in average effect sizes might also be explained by the scope of what each valuation method measures. As discussed in Chapter 1, contingent valuations can include nonuse values, while the hedonic pricing and travel cost methods only capture use values. Thus, one can reasonably expect contingent valuations to have larger ranges of responses. But even when the three methods are measuring the same scope of value of an environmental good, there are theoretical reasons why the results from different methods should differ. As discussed in Chapter 1, the travel cost method and hedonic pricing are properly interpreted as lower bounds of maximum willingness to pay. Thus, they should be less than or equal to the maximum willingness to pay elicited from contingent valuation. However, others have shown that results from contingent valuation should be smaller due to the differences between Hicksian and Marshallian demands measured by contingent valuation and the weak complementarity methods, respectively (Smith & Pattanayak, 2002). The results from this meta-analysis clearly support this latter explanation.

Thus, the disparity between effect sizes of contingent valuation studies and those from the hedonic pricing and travel cost methods can be attributed to each aspect of the contingent valuation method: the researchers' designs, the respondents' replies, and the scope of the measurement instrument. As Randall, Hoehn, and Brookshire (1983) have noted, contingent valuations result in highly contingent results.

How are these results explained vis-à-vis other studies that demonstrate convergent validity? Those studies do not compare the results of different studies. Rather, they are studies that employ two or more valuations to the same case (e.g., Farber 1988; Walsh, Ward, & Olienyk, 1989). While such experiments have superior controls for spurious effects – such as environmental good, change in good, and type of description of good – their results only demonstrate the comparability of employing the methods under carefully controlled, static conditions. They speak nothing on the reliability of the results in varying conditions. Such information is critical to the selection of methods for original studies and the use of existing results in benefit transfers. The results presented in this research address this information need. Thus far, across the broad range of applications, the effect sizes of contingent valuation and the other two methods are not convergent. The conditions under which convergence is stronger and weaker are addressed later in this chapter.

Moderating Effects of Other Variables of Interest

So far, all four primary independent variables have been shown to moderate environmental valuations. Before their relative effects were analyzed together, a few more variables of interest were considered for their moderating effects. First was the income of subjects. Income is well understood to limit and affect the choices consumers make in the market, and a meta-analysis of it could measure its moderating effect on environmental valuation. However, only 29 studies in the data set (12.7%) reported *any* measure of income, and those measures varied by monetary denomination, year, and scope (e.g., household versus personal income). Therefore a meaningful analysis of this variable was not possible. The lack of reported incomes is disappointing, given the theorized effect of income on valuation. Through the U.S. Census Bureau it may be possible to gather comparable income statistics of U.S. samples in the studies, knowing their locations and years of study. Doing so would enable the measurement of the correlation between income and effect size. Such an effort is beyond the scope of this project, but reserved for future extensions of this work.

Second was the years of data collection. While the years of publication were used as a proxy for the refinements of the valuation methods, the years of data collection were used to measure temporal changes in values for the environment[16]. That is, analyzing subjects' valuations over time could reveal any broad changes in preferences for the environment. One study has shown, for example, that people's attitudes toward water conservation changes with recent experiences with drought or flooding (Trumbo, Markee, O'Keefe, & Park, 1999).

In many cases in this study, the data were collected over more than one year. In these cases, the mid-points of the years were assigned. Pearson's correlation coefficient[17] was calculated between these midpoint data collection years and the effect sizes, weighted by the sample sizes of the observations, resulting in a statistic of $r = -0.11$[18]. The value is statistically significant, even though the strength of the relationship is quite

---

[16] In only one case in this data set were the years of data collection and publication the same. The difference between these years in all the cases ranged from 0 to 25 years.

[17] Pearson's correlation coefficient, r, estimates the degree of linear relationship between two ratio or interval variables. Its values range from $-1 \leq r \leq +1$, where -1 signifies a perfect inverse relationship, 0 signifies no relationship, and +1 signifies a perfect positive relationship.

[18] In this case, the total sample size was T = 743,084 (356 observations weighted by their respective sample sizes). At $\alpha = 0.01$, the critical value of r is about 0.08. The measured value is greater than the critical value.

mild. It indicates that among all the observations included in this data set, effect size is mildly negatively associated with year of data collection. That is, there has been a slight decline in effect sizes over the years. Of course, this analysis spans all the other variables already shown to be significant moderators of effect sizes. It is possible that the relationship is stronger, or even positive, for specific environmental goods or under other specific circumstances. This variable was included in the meta-regression analyses reported later in this chapter.

Third was the location of the study, to determine whether effect sizes vary by region. This analysis could shed light on the geographic generalizability of valuation results. One study has already shown that interstate transfers (within the United States) are problematic, even when the same measurement instrument is used, while intrastate transfers can work well (Boyle, Bergstrom, & Poe, 2001). That is, location moderates value.

As described in Chapter 4, the distribution of cases across regions is heavily skewed toward the United States. Within the United States the cases are concentrated in a few western and northeastern states, with a notable lack of studies of southern states. Two meta-analyses were conducted to test the influence of region on effect sizes. First, the cases were split between those conducted in the United States and all other countries. Table 5.9 summarizes the results. Of the 416 usable cases, 349 were in the United States, while only 67 were elsewhere. The average effect size in the United States is 7.54 standard deviations, while that for all other countries is 3.29. A t-test of the difference between these two means resulted in a test statistic of $t = 5.17$: the difference is statistically significant[19], with effect sizes from the United States averaging more than twice those from other countries.

Table 5.9: Mean Effect Sizes by Country

	sample size T	**# of effect sizes K**	**mean effect size D**	**Var(d)**	**Var(e)**	**Var(e) / Var(d)**	**χ^2 test statistic Q**	**99% C.I. for δ**
Combined	769558	416	6.94	109.92	0.02	0.01%	3005383	5.60 8.29
U.S.	662177	349	7.54	124.29	0.02	0.01%	2537442	5.98 9.10
All others	107381	67	3.29	21.25	0.01	0.03%	241833	0.56 1.80

Next, as a cursory analysis of the moderating effect of intranational regions, the cases within the United States were divided by state. While 38 states are represented in the data set, 27 of them have fewer than 10 cases. Thus, this analysis was limited to the 10 states with the most observations. They are California, Colorado, the District of Columbia, Illinois, Massachusetts, Maine, Oregon, Pennsylvania, Wisconsin, and Wyoming. Table 5.10 summarizes the results. Combined, the average effect size from the 10 states is 1.11 standard deviations. Separated, the effect sizes range from 0.53 in the District of Columbia to 5.81 in California. A one-way ANOVA resulted in a test statistic of F = 9.30, confirming that not all of the means are equal[20]. Fisher's LSD test

[19] A t-test of the difference between two means, with unequal variances, resulted in a test statistic of t = 5.17. At $\alpha = 0.01$, and 66 degrees of freedom, the critical value is t = 2.6. Thus, with 99% confidence the null hypothesis of equal means is rejected.

[20] At $\alpha = 0.01$, with 9 degrees of freedom in the numerator and 242 degrees of freedom in the denominator, the critical value is F = 2.5. The test statistic is larger. Thus, with 99% confidence the null hypothesis of the equality of all means is rejected.

was again employed for the subsequent pair-wise analyses. Table 5.11 shows the t-statistics from the test, highlighting those values that indicate a significant difference between the values of the pairs of states. Interestingly, California is the only state of these 10 that has a significantly different average effect size than others. In fact, its average effect size is significantly larger than each of the other states analyzed except Wisconsin.

The results of these regional analyses seem to support two popular perceptions. First, while the United States has been criticized as leading the way in resource depletion and environmental destruction, it has also been seen as a leader in environmental protection. This is a sweeping statement, but Table 5.9 supports the claim. Values Americans impart on improvements to the environment are, on average, greater than those of the other countries in this data set. Second, within the United States there is a popular perception of California as a more politically liberal, “left coast” state with stronger support for environmentalism. Again, this statement sweeps over the widely varied political terrain within the state, and all the confounding variables that might influence this finding, but the results shown in Table 5.11 certainly support the perception. They also suggest, contrary to another study (Boyle, Bergstrom, & Poe, 2001), that benefit transfers across state lines can be reliable except when one of the states is California.

Table 5.10: Mean Effect Sizes by U.S. States

	sample size T	# of effect sizes K	mean effect size D	Var(d)	Var(e)	Var(e) / Var(d)	χ^2 test statistic Q	99% C.I. for δ
Combined*	369864	252	1.11	39.59	0.00	0.01%	3166006	0.08 2.14
CA	9401	45	5.81	51.22	0.10	0.20%	22867	2.79 8.82
CO	8966	35	0.78	29.05	0.02	0.06%	60015	-1.72 3.29
DC	6100	20	0.53	30.18	0.01	0.05%	44171	-2.98 4.04
IL	11521	28	1.09	29.99	0.01	0.04%	74853	-1.78 3.96
ME	3436	21	1.77	18.25	0.03	0.19%	11133	-0.88 4.42
MA	4658	16	1.80	17.29	0.02	0.11%	14249	-1.27 4.86
OR	306102	21	0.95	41.60	0.00	0.00%	2859793	-3.05 4.96
PA	10738	23	1.07	23.52	0.01	0.04%	54929	-1.78 3.92
WI	3558	11	3.57	27.50	0.03	0.12%	9384	-1.48 8.61
WY	5384	32	0.79	26.59	0.03	0.10%	32826	-1.72 3.29

* Includes only the 10 states in this table, not all 38 represented in the data set.

Table 5.11: t-Statistics from Fisher's LSD Test of the Differences Between the Mean Effect Sizes of Pairs of States

	CA	CO	DC	IL	MA	ME	OR	PA	WI	WY
CA	---									
CO	6.18*	---								
DC	5.44*	0.25	---							
IL	5.43*	-0.34	-0.53	---						
MA	3.82*	-0.93	-1.05	-0.63	---					
ME	4.24*	-0.99	-1.10	-0.65	0.02	---				
OR	5.09*	-0.17	-0.37	0.13	0.71	0.73	---			
PA	5.12*	-0.30	-0.49	0.02	0.62	0.64	-0.11	---		
WI	1.85	-2.23	-2.24	-1.93	-1.25	-1.34	-1.95	-1.88	---	
WY	6.02*	-0.01	-0.25	0.33	0.91	0.97	0.16	0.29	2.21	---

* Significantly different at $\alpha = 0.01$

Combined Effects of Moderators: Meta-Regressions

With several moderating variables identified, the analysis continued with evaluations of their relative effects on environmental valuation. Specifically, regression analysis was used to evaluate the variance in effect sizes due to the moderating variables together[21]. Table 5.12 summarizes three models that cover all the moderating variables in this study. In all of these models, observed effect sizes were weighted by their respective sample sizes, as has been throughout this chapter.

The independent variables in the models include four of the six analyzed environmental goods: air, animal, land, and water. There were too few observations of plant, toxic, and other goods for statistical analysis. (See Table 5.1.) The four remaining

[21] Regression analyses result in linear, relational equations of the general form

$$Y = b_0 + b_1X_1 + b_2X_2 + \ldots + b_kX_k + e,$$

where Y is the dependent variable, X_i are the independent (explanatory) variables, b_i are the coefficients of X_i (i.e., the slope of the line between X_i and Y), b_0 is the y-intercept (the value of Y when all $X_i = 0$), and e is an error term. In this case, Y is effect size, and X_i are the moderating variables. Any b_j that is found to be statistically significant describes the change in effect size for a unit change in X_j, while holding all other X_i constant. Thus, the relative effects of the moderating variables were assessed.

environmental goods were coded as three dichotomous variables. Specifically, “Air,” “Animal,” and “Land” were coded as “yes” or “no,” and water was the reference good signified by negative codes for all of the other three goods.

Furthermore, the method of valuation was collapsed into one dichotomous variable for contingent valuation (“CV”). In effect, this combines the hedonic pricing and travel cost methods into one reference group. This is justified for several reasons. On a theoretical level, both methods are based upon real market behavior surrounding weak complements of the environmental goods. Also, many of the travel cost analyses in this data set measured environmental values by having environmental measures as independent variables in the trip generation functions[22]. That is, they were hedonic pricing models of trip generation. Finally, on a practical level, there are too few travel cost valuations to make it a statistically significant category by itself. Thus, method of valuation was coded as a dichotomous variable for contingent valuation, with weak complementarity methods as the reference method.

Similarly, the location of study was converted to two dichotomous variables. Country was coded as United States or other (“U.S.”), and state was coded as California or other (“CA”). These categories were already shown to moderate effect size. (See Tables 5.9, 5.10, and 5.11.) The magnitude of change (“Holistic change”), type of description of change (“Quantitative description”), and era of contingent valuation development (“CV>1993”) were already operationalized as dichotomous variables, and they were included in the regression analyses as such. Finally, the midpoint years of data collection (“Data year”) are ratio data and were included in the regression in that form.

[22] This might also explain why the travel cost and hedonic pricing methods were not differentiated as moderating variables. (See Tables 5.5 and 5.6.)

Table 5.12: Comprehensive Models of Effect Size

	Model 1: Full set	Model 2: CV subset	Model 3: U.S. subset
total sample size, T	708916	31235	637738
# effect sizes, K	325	121	283
adjusted R^2	0.56	0.12	0.57
F	113081.9*	539.9*	106274.5*
Independent variables	**b** *t statistic*	**b** *t statistic*	**b** *t statistic*
y – intercept	-923.95	951.00	-1053.12
	*-195.70**	*30.20**	*-198.47**
Air	12.29	-4.68	11.96
	*394.53**	*-27.19**	*362.58**
Animal	-1.19	-0.57	-3.11
	*-13.31**	*-4.39**	*-33.14**
Land	-6.85	-0.15	-8.22
	*-227.76**	*-1.57*	*-250.20**
Holistic change	-2.51	2.53	0.97
	*-47.00**	*16.83**	*10.63**
Quantitative description	-0.98	-1.16	1.46
	*-18.66**	*-11.94**	*16.47**
CV	-1.91		-1.07
	*-39.71**		*-17.56**
CV > 1993		5.83	
		*47.44**	
U.S.	-3.98	2.71	
	*-121.64**	*27.89**	
CA			-3.57
			*-42.01**
Data year	0.47	-0.48	0.53
	*197.99**	*-30.30**	*199.88**

Dependent variable = effect size

* Significant at $\alpha = 0.01$ and T-1 degrees of freedom

Model 1 includes all valuations of the four environmental goods. The reference group[23] is valuations of incremental, qualitative changes in water, conducted with a weak complementarity method in a foreign country. The F statistic indicates that the model is significant[24], and the adjusted R^2 indicates that 56% of the variance in effect size is explained by the model[25]. Indeed all of the independent variables are statistically significant[26]. Most of the b coefficients in the model have signs (positive or negative) that are consistent with previous findings in this chapter. For example, improvements to air produces greater effect sizes than those of water, all other things being constant. Similarly, animal and land goods produce smaller effect sizes than water. Also, holistic changes produce smaller effect sizes than incremental changes. Most importantly for this research, the model confirms the previous finding that contingent valuations produce smaller effect sizes than the travel cost and hedonic pricing methods. Two independent variables have the opposite signs of what was expected. In Model 1, quantitative

---

[23] In a regression equation that includes dichotomous independent variables, the reference group characterizes observations when all the dichotomous variables are "no."

[24] Analysis of variance of a regression equation tests the overall significance of the model, and results in a F statistic. Formally, the null and alternative hypotheses are

H_0: $b_1 = b_2 = \ldots = b_k = 0$ (the model is insignificant)

H_1: not all $b_k = 0$ (the model is significant).

The null hypothesis is rejected if the test statistic is greater than the critical value (defined by two separate degrees of freedom and a selected level of significance, α), otherwise the null cannot be rejected. Rejecting the null indicates that the model is significant.

[25] While the adjusted R^2 quantifies the explained variance *in this model*, it overestimates the explained variance in the represented relationship, because the variance in the original studies' observations are not accounted in this model.

[26] The significance of each independent variable was determined with a t-test of the value of the coefficient, b. Formally, the null and the alternative hypotheses are

H_0: $b = 0$ (the independent variable is not related to the dependent variable)

H_1: $b \neq 0$ (the independent variable is related to the dependent variable).

The null hypothesis is rejected if the test statistic lies outside the critical values (defined by the degrees of freedom in the sample and the selected level of significance, α), otherwise the null cannot be rejected. Rejecting the null indicates the independent variable is related to the dependent variable, and the value of b measures the relationship.

descriptions of change produced smaller effect sizes than qualitative ones, contrary to earlier findings. And valuations conducted in the U.S. produce smaller effect sizes than those done abroad. These will be discussed when comparing them with their effects in Models 2 and 3.

Model 2 includes only the subset of valuations using contingent valuation, and its purpose was to confirm and measure the relative effect of the refinement of contingent valuation ("CV>1993"). Its reference group is valuations of incremental, qualitative changes in water, conducted through 1993 in a foreign country. Its F statistic indicates the model is significant, and all but one independent variable ("Land") were found to be significant. The model does confirm the moderating effect of the development of the contingent valuation method up through 1993. The coefficient indicates that, holding all other variables constant, contingent valuations published after 1993 produced effect sizes that average 5.83 standard deviations greater than those published up through 1993. The difference is larger than that reported in Table 5.8.

Model 3 includes only the subset of valuations conducted in the United States, and its purpose was to test the effect of California ("CA") on effect sizes. Its reference group is valuations of incremental, qualitative changes in water, conducted with a weak complementarity method outside of California. Like Model 1, all of the variables in Model 3 are statistically significant. The model does confirm the moderating effect of studies done in California, but the sign of the coefficient contradicts the previous finding. Tables 5.10 and 5.11 show California producing larger effect sizes, while Model 3 shows it producing smaller ones. The contradiction was explored further by comparing the effect for different environmental goods. These results are reported later in this chapter.

Considering all three models together the primary variable of interest, method of valuation, confirmed earlier findings: contingent valuation produces smaller effect sizes than hedonic pricing and travel cost methods, but the gap – while still significant – has lessened in the last decade. Also, not surprisingly, the type of good moderates value in all three models. However, the magnitude of change – holistic versus incremental – produced varying results. While holistic changes produced smaller effect sizes in Model 1, it produced larger ones in Models 2 and 3. This might reflect genuine differences in its effects in the different models (i.e., in contingent valuations and in the U.S. holistic changes produce larger effect sizes), and one could rationalize why this would be so. But another explanation that cannot be overlooked is the validity of the measurement. As discussed in Chapter 4, the two categories of magnitude are not completely mutually exclusive. What is considered an incremental change in one study may be a holistic one in another, even though the actual change in each is the same. Also, the two categories wash over the wide ranging degrees of changes in the studies. While the operational categories were necessary for analytic purposes, their utility is questionable. A better measure of magnitude of change is needed: one that can be validated so its influence on effect size can be more confidently assessed. The type of description of change – quantitative versus qualitative – also produced varying results. While the earlier analysis showed that quantitative descriptions result in larger effect sizes, Models 1 through 3 reveal mixed conclusions. This may be due to the variable's association with magnitude of change. As shown in Table 4.4, quantitative descriptions are strongly associated with incremental changes, and qualitative descriptions are associated with holistic changes. Thus, the mixed results in Models 1 through 3 may reflect the multicollinearity between these two variables[27]. "U.S." also produced mixed results. In Model 1, the resulting relationship is opposite that revealed in Table 5.9. One possible explanation may be the relatively few cases done outside the United States, making its effects more volatile to changes in the subsets. Finally, the years of data collection also produced mild, mixed results. This reflects the weak relationship described earlier in its correlation coefficient. Indeed, a scatter plot of effect sizes by year of data collection, Figure 5.1, shows widening effects over time, but no strong trend upward or downward.

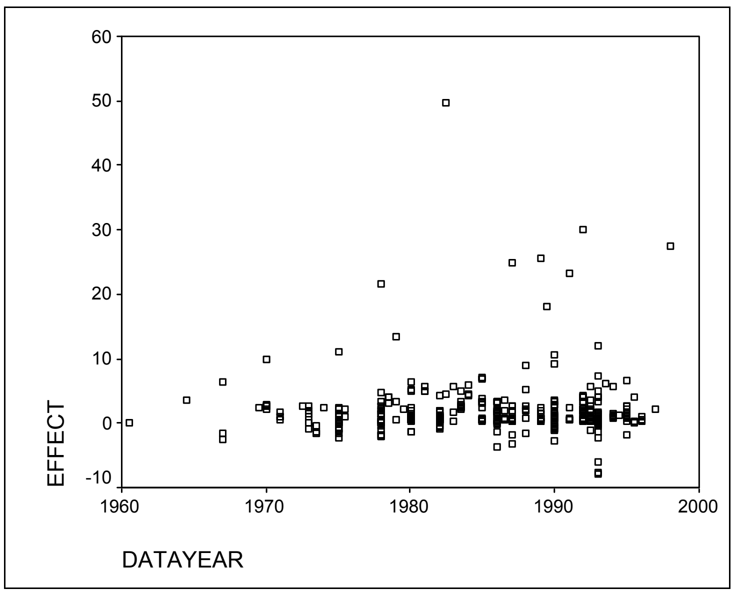


Figure 5.1: Effect Size by Year of Data Collection

Subsets by Environmental Goods

To further explore the effects of the method of valuation, and the contexts in which the effects change, the data set was divided into the four primary environmental goods and regression equations were recalculated with the remaining primary

[27] However, when Models 1 through 3 were recalculated without "Holistic change" in the models, "Quantitative descriptions" still produced mixed results.

independent variables[28]. Table 5.13 summarizes the results and shows the varying effects of method of valuation among the environmental goods. While contingent valuation was shown to produce lower effect sizes overall, as confirmed in Model 1, it produces larger effect sizes for enhancements air and land. For air, contingent valuation produces effect sizes 0.68 standard deviations larger than those from the hedonic pricing and travel cost methods, while holding the other variables constant. For land, the increase is 2.44 standard deviations. For animal and water goods, contingent valuation produces lower effect sizes, consistent with previous findings. Surprisingly, the effects do not appear to be associated with experience of using a specific method with each good (see Table 4.6). Contingent valuation has been disproportionately used to value animals, disproportionately underused to value air, and roughly proportionally used to value land and water. The effects of contingent valuation do not match its use.

The effects of contingent valuation's refinements were also explored by recalculating Models 4 through 7 with a data set limited to only contingent valuations, and substituting year of publication ("CV>1993") for method of valuation. Table 5.14 summarizes the results. For air and land, measures of effect size by contingent valuation are significantly larger after 1993, consistent with findings in Tables 5.8 and 5.12. For animal and water goods, effect sizes from contingent valuation have changed slightly or insignificantly over the years.

[28] Years of data collection ("Data year") was excluded from these models because of its weak relationship with effect size, and country ("U.S.") was excluded because of its low variance.

Table 5.13: Models of Effect Size by Environmental Good

	Model 4: Air	Model 5: Animal	Model 6: Land	Model 7: Water
total sample size, T	217189	16028	417166	87612
# effect sizes, K	135	67	110	102
adjusted R^2	0.06	0.77	0.04	0.03
F	4835.2*	17564.6*	5972.2*	782.6*
Independent variables	**b** *t statistic*	**b** *t statistic*	**b** *t statistic*	**b** *t statistic*
y – intercept	7.68	4.99	3.29	5.57
	*27.13**	*208.23**	*106.76**	*26.41**
Holistic change	-8.91	0.74	-2.84	-0.50
	*-31.75**	*9.54**	*-95.65**	*-2.20***
Quantitative description	11.48	0.10	-1.46	2.37
	*40.64**	*6.70**	*-48.77**	*11.36**
CV	0.68	-5.10	2.44	-2.95
	*4.71**	*-67.84**	*72.62**	*-18.48**

Dependent variable = effect size
* Significant at α = 0.01 and T-1 degrees of freedom
** Significant at α = 0.05 and T-1 degrees of freedom

Table 5.14: Models of Effect Size by Environmental Good - CV Subset

	Model 8: Air	Model 9: Animal	Model 10: Land	Model 11: Water
total sample size, T	6993	13533	16170	19567
# effect sizes, K	47	62	49	43
adjusted R^2	0.77	0.01	0.14	0.01
F	7831.98*	42.32*	882.86*	86.20*
Independent variables	**b** *t statistic*	**b** *t statistic*	**b** *t statistic*	**b** *t statistic*
y - intercept	-1.70	0.66	-1.31	1.27
	*-4.33**	*57.14**	*-7.56**	*6.72**
Holistic change	-4.46	NA	2.67	2.11
	*-19.39**		*18.32**	*15.73**
Quantitative description	14.96	0.10	-0.95	0.22
	*71.57**	*7.04**	*-9.82**	*1.47*
CV>1993	7.00	-0.10	3.04	-0.28
	*28.57**	*-5.66**	*33.20**	*-1.92*

Dependent variable = effect size
NA = not applicable; there is no variance in this variable
* Significant at α = 0.01 and T-1 degrees of freedom

Finally, the effect of California was reexamined, in light of the apparent contradictions between Tables 5.11 and 5.12. The data set was limited to studies conducted in the United States, and the four regression equations were recalculated with a dichotomous variable for California ("CA"). Table 5.15 summarizes the results. For three of the four environmental goods – air, animal, and land – California is a significant variable. Valuations of air conducted in California result in substantially smaller effect sizes than elsewhere, holding the other variables constant. On the other hand, valuations of land conducted in California result in substantially larger effect sizes. Valuations of animals result in slight differences between California and other states, and valuations of water result in no significant difference. These findings appear to explain the

Table 5.15: Models of Effect Size by Environmental Good – U.S. Subset

	Model 12: Air	**Model 13: Animal**	**Model 14: Land**	**Model 15: Water**
total sample size, T	212314	15108	347657	77711
# effect sizes, K	126	61	72	81
adjusted R^2	0.09	0.78	0.04	0.02
F	5237.2*	13510.0*	3627.6*	314.1*
Independent variables	**b**	**b**	**b**	**b**
	t statistic	***t statistic***	***t statistic***	***t statistic***
y - intercept	23.95	5.05	3.01	6.99
	*55.78**	*207.97**	*40.96**	*26.41**
Holistic change	-3.75	0.68	-2.68	-0.13
	*-7.91**	*8.80**	*-36.03**	*-0.43*
Quantitative description	-4.48	0.04	-1.76	1.21
	*-10.45**	*2.54***	*-23.96**	*4.62**
CV	-10.17	-5.08	1.19	-3.67
	*-17.64**	*-68.22**	*19.14**	*-16.62**
CA	-13.14	-0.39	11.69	-0.49
	*-110.68**	*-9.30**	*89.56**	*-1.38*

Dependent variable = effect size
* Significant at $\alpha = 0.01$ and T-1 degrees of freedom
** Significant at $\alpha = 0.05$ and T-1 degrees of freedom

contradiction between the previous findings. Valuations in California do indeed result in different effect sizes, but the differences are dependent upon the environmental good.

Limitations of the Results

Before the implications of these results are discussed in Chapter 6, it is important to recognize their limitations. The most apparent is the unexplained variance in the regression models. While each model is statistically significant, most of them leave a vast majority of the variance in effect sizes unexplained. One obvious reason for this is the bluntness of the operational measures of the independent variables. The environmental good categories (except air), magnitudes of change, and even methods of valuation are much more varied than the few categories used to measure them: Table 4.1 presents dozens of environmental goods reduced to seven categories; the magnitudes of change span a continuum, yet were reduced to two categories; and the variants of each method of valuation are hidden in the three categories used to capture them[29]. However, given that the purpose of this research is to assess the broad (meta-level) effects of these variables on environmental valuation, the broad categories are justified, and the large, unexplained variance is expected and accepted as a consequence of the approach. Future research could focus on the effects of each variable and use more sensitive measures. Also, the unexplained variance might be caused by other moderators not included in this research. All of the meta-analyses conducted in this research suggest the presence of other moderators.

[29] The methods were further reduced to two categories (contingent valuation and weak complement methods)!

Another limitation of the results is the differing subsets of studies used for different analyses. Different cases had to be eliminated from different analyses because of the differences in missing data in the observations. The differing subsets present a possible spurious effect in the analyses. That is, the results presented may be due to the differing subsets of data rather than the variables observed. While the magnitude of such a spurious effect is unknown, it is minimized by the large sample size in this meta-analysis. One clear conclusion learned from this, however, is that information reported in environmental valuation studies is far from standard. Improved and standardized reporting in original valuation studies is necessary to increase the validity of meta-analyses such as this, and to make benefit transfer more viable (Boyle & Bergstrom, 1992).

A common limitation of any meta-analysis is the possibility of publication bias (Begg, 1994). This is a bias in the meta-analysis data set resulting from the presumed biases of journals to publish significant results. If journals tend to publish significant results, and not publish insignificant results, then a meta-analysis of articles would result in an overstated effect size. This problem is not of great concern to this meta-analysis for two reasons. First, the problem is most relevant to meta-analyses focused on synthesizing and measuring the *absolute magnitudes* of effects. The purpose of this meta-analysis, however, is to measure the *relative magnitudes* of effects from different valuation methods. That is, the primary question in this research is one of convergence or divergence, not magnitude. Thus, even if a publication bias was significant, it would have to be different for the different valuation methods to have an effect on these results. For example, consider the difference in means tests reported in Table 5.6, in which the average effect sizes from contingent valuations and hedonic pricing studies were shown to be statistically different. A "file drawer analysis"[30] (Hunter & Schmidt, 1990, pp. 512-513) estimated the number of unpublished hedonic pricing studies with an average effect size of zero that would have to exist in order to eliminate the difference between the average effect sizes of contingent valuation and hedonic pricing studies:

$$X_{HP} = k[(d_{k\text{-average}}/d_{c\text{-average}}) - 1] = 204[(2.25/1.62) - 1)] = 80.14,$$

where k is the number of hedonic pricing studies, $d_{k\text{-average}}$ is the unweighted average effect size of hedonic pricing valuations, and $d_{c\text{-average}}$ is the unweighted average effect size of contingent valuations. The results estimate that 80 unpublished, insignificant hedonic pricing studies would have to exist to change the conclusion of Table 5.6. Given the large sample of studies used in this evaluation – larger than any other meta-analysis reviewed in this research – a comparably large number of other studies would be needed to negate the findings[31]. The second reason why publication bias is not of great concern to this research is that many of the studies in the data set did indeed report insignificant results of environmental valuations.

A final limitation of these results is perhaps the most important: its generalizability. The sample of studies in this research is not representative. Rather, they were deliberately selected from the most prolific journals in environmental valuation to

[30] The name of the analytic procedure refers to the studies with insignificant findings that occupy the file drawers of researchers' cabinets instead of the pages of journals, because of a publication bias.
[31] For comprehensiveness, the "file drawer analysis" was repeated for the absolute magnitude of environmental valuations. Considering the overall average effect size of the entire data set, 251 unpublished valuations averaging zero-effect would have to exist to make the overall average effect size insignificantly different from zero (at a $\alpha = 0.01$ level of significance).

presumably capture high quality studies. This strategy was used to qualitatively minimize the chances of a type 1 error: rejecting the convergence of the valuation methods when they are actually equal. The assumption was that the most prolific valuation journals publish higher quality valuation studies, and such studies would best represent the practices of environmental valuation in a meta-analysis. Thus, the likelihood of erroneously rejecting convergence would be less than that for a data set that included lower quality studies. The consequence of this strategy is that the magnitude of the divergence is not representative. The advantage, however, is that the conclusion of divergence is better substantiated.

Summary

In summary, the average effect size of all the environmental valuations in this data set is 6.90 standard deviations. However, the variance in this average strongly suggests the presence of moderating variables. Several variables were tested individually and collectively for their moderating effects. The conclusions of the analyses are as follows:

- Type of good is a significant moderator. Air and water have significantly larger average effect sizes than land and animal, and air has a significantly larger average effect size than water.

- Magnitude of change and type of description of change are significant moderators, but their effects are mixed. These conflicting results may be due to validity problems in the measures themselves.

- The method of valuation, the primary variable in this research, is a significant moderator. Effect sizes from contingent valuations are, on average, significantly smaller than those from the travel cost and hedonic pricing methods. There is no significant difference between the effect sizes of the travel cost and hedonic pricing methods. On average, contingent valuation produces effect sizes that average 40% to 55% less than those of the hedonic pricing and travel cost methods. The difference varies with the environmental goods valued. While contingent valuations of animals and water result in lower effect sizes, those of air and land result in larger effect sizes. There is evidence, however, that contingent valuations have become more reliable over time. Those published after the 1993 National Oceanic and Atmospheric Administration report (Arrow et al., 1993) are less variant than those published before it; the difference is accentuated for valuations of air and land. Even so, the difference between the results of contingent valuations and travel cost or hedonic pricing methods remain significant.

- Few studies report statistics on the incomes of the subjects, preventing an analysis of the moderating effect of income on valuations.

- The year of data collection is not a significant modifier. There is not significant evidence to demonstrate changes in environmental values over the years.

- Location of study is a significant moderator. Effect sizes in the United States average twice the magnitude of all other countries in the data set, but the relatively small number of foreign studies in the data set casts doubt on this finding. Within the United States effect sizes in California are, on average, significantly different than those from other states in the data set. Specifically, effect sizes for air tend to be much smaller in California, and those for land tend to be much larger.

CHAPTER 6

IMPLICATIONS

This research set out to evaluate three popular methods of environmental valuation, because their substantial impact on public policy warrants a critical analysis, and because they have been clouded by controversy. The evidence presented sheds light on the methods so policy practitioners can better understand their results and judge their proper use. This chapter folds the results of this research back into the broader debate over environmental valuation to interpret the results and their implications in environmental policy. It begins by describing what the results of environmental valuations tell us. This includes a discussion of the convergent validity of the methods and their defensible uses. It then examines the literature to describe what environmental valuations do not tell us, but should.

What Environmental Valuations Tell Us

Following the analyses presented in Chapters 4 and 5, the most fundamental question that must now be asked is, "What do environmental valuations tell us?" The evidence in this research does not completely support one side of the debate described in Chapter 2, but it certainly leans with the critics. Interestingly, the evidence brings attention to a second question that appears equally important: "*Whose* values do they represent?"

Addressing the first question, proponents and opponents agree that environmental valuations are intended to measure economic values of environmental goods. This dimension of value has several components of its own, which are illustrated in Figure 6.1. Total economic value of environmental goods is composed of use values and nonuse values, and these two categories are composed of five subcategories of value: direct use, indirect use, option, existence, and bequest values[1] (Barbier, 1994). But the three methods of valuation do not all claim to measure total economic value, only portions of it. The hedonic pricing and travel cost methods, each being based upon observed market behavior, can only measure direct use values. Contingent valuation, in contrast, is claimed to be able to measure all types of economic value, but these distinctions are often absent in its applications and left to the judgments of its readers. Thus, the first

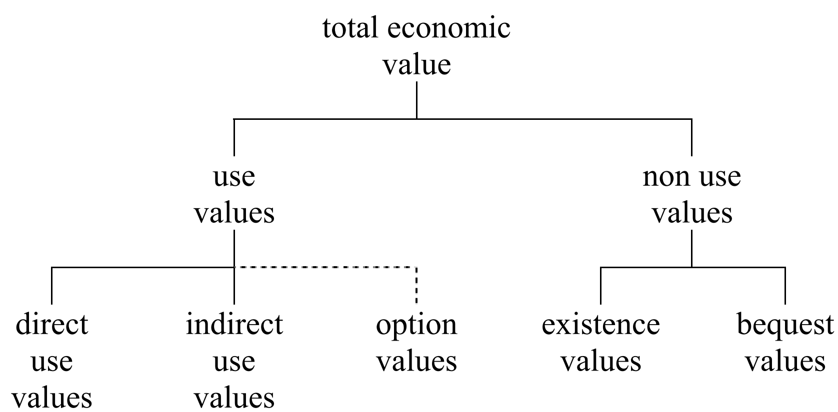


Figure 6.1: Economic Values of Environmental Goods (adapted from Barbier, 1994)

[1] Direct use values are those derived from direct consumption or interaction with the good. Indirect use values reflect benefits provided by environmental goods to support economic life, such as ecological services. Option values do not represent a separate category of values, but reflect future use values: the value of holding the option for future use. Existence values are intrinsic to the good and unrelated to the use of the good by the valuer or anyone. Bequest values reflect desires to conserve the environment for use by future generations.

conclusion that can be made is that environmental valuations speak of direct use economic values, and in some cases (contingent valuation) it may also speak of other dimensions of economic value.

How authoritatively do they speak of these values? The evidence from this study is mixed. It clearly says that environmental values are positive. In all the subsets of cases observed, the average effect size is always positive. In only a few cases do the confidence intervals for the average effect size include zero, and each of those cases have a relatively small number of observations[2]. The consistency with which positive measures of values were attributed to environmental improvements lends some credibility to the methods. That is, the methods measure the *direction* of value well[3]. But what can be said about the *magnitudes*? Do they accurately express the same economic values of environmental goods? The divergence of effect sizes among the methods say no. Contingent valuations and the two weak complementarity methods produce significantly different average effect sizes, while controlling for environmental good, holistic versus incremental changes, and quantitative versus qualitative measures of change. The divergence suggests that one of the methods has problems with validity, or is measuring different aspects of value, as further discussed below. In lieu of true economic value, others have offered alternative interpretations of the magnitudes of results. Nunes and van den Bergh (2001) note that the range of environmental benefits stemming from an environmental good often extends far beyond that which is described or inferred in a valuation study. As a result, the estimate of value can at best be taken as an incomplete,

---

[2] Smaller numbers of observations increase the widths of confidence intervals.

[3] This claim is based on the assumption that values for environmental improvements are indeed positive, which is substantiated by opinion surveys on environmental values (e.g. Kempton, Boster, & Hartley, 1999).

lower bound of economic value. Gregory, Lichtenstein, Brown, Peterson, and Slovic (1995) found that respondents' stated values are only vaguely represented in monetary terms. Any dollar amount within a broad range is considered an acceptable expression of value, but the magnitudes are imprecise and not appropriate for economic analyses:

> "The more fundamental question raised by our results is whether holistic monetary estimates, as required by [contingent valuation] elicitations, provide appropriate measures of value for many nonmarket environmental goods. Our conclusion . . . is that often they do not. The imprecision in monetary estimates reported here arises because statements of monetary value for these goods are constructed by participants in the course of the elicitation process." (p. 471)

At first glance, the answer to the second question seems simple: the values expressed or revealed are those of the subjects, be they respondents in a contingent valuation study, home buyers in a hedonic pricing study, or visitors in a travel cost study. But this interpretation belies the contextual problem of the valuation methods. In Chapter 5 this problem was identified as information bias: how the results of contingent valuations are significantly influenced by the information provided by the researchers. A similar problem exists for the hedonic pricing and travel cost methods. The selections of goods to include in the travel cost and hedonic pricing models are, in essence, researchers' judgments of the weak complementary bundles of goods consumers are purchasing with their properties and travels. The reality of what consumers are consciously purchasing may range from none of the environmental goods identified by the researchers to entire spectra of economic values of the goods. For all three methods, researchers frame the environmental goods being valued by identifying the goods, specifying their scopes, and selecting the measures of them. None of these tasks is an objective decision, and all of them influence or define the results of valuation studies. Thus, when assessing the meanings of the outputs of environmental valuations, it is important to recognize that the values expressed are not only those of the subjects, but also those of the researchers (Nunes & van den Bergh, 2001).

One value of researchers that is consistently expressed in the applications of these methods is in the scope of value itself. Just as researchers frame the environmental good being valued, so too do they frame how value is measured. For each of these three methods, value is limited to the economic view and it is measured in monetary terms (Kassiola, 2003a). This framework is defended by proponents of environmental valuation. The economic view is justified, as described in Chapter 2, because it has become the dominant perspective in public policy analysis (Rees, 1994). Therefore the economic value of environmental goods must be quantified so they are not excluded from cost-benefit analyses (Goodland & Ledec, 1994). The use of monetary metrics to measure economic value is strictly a pragmatic decision. It is not so much justified by proponents as it is irrelevant to them:

> "Opponents of economic evaluation frequently misrepresent economics as being about money: it is not, money is only the yardstick with which to compare the relative values of different goods. . . What the economist seeks to do is derive a rigorous method of measuring the values different individuals place upon different goods in order that these values can be compared." (Green & Tunstall, 1991, p. 125)

To practitioners of the methods, money is just a numeric measure that is convenient because of its well understood, common meaning. Any other scaled measure would be welcome, but none is as commensurate among different people.

However, the selection of money as the measure *is* relevant because the common meaning it holds is limited in scope. Money is the currency of markets, expressing economic values. But there is significant evidence that people object to expressing their values for environmental goods in solely economic terms, and find it difficult to do so. Sagoff (1998) argues that individuals' policy judgments expressed in contingent valuations are not actually their preferences as utility maximizers but normative expressions of what society ought to do as a collective. Empirical evidence appears to support this claim. Kahneman and Knetsch (1992) tested the construct validity of contingent valuations and found that results actually represent the respondents' willingness to pay for the moral satisfaction of contributing to the public goods, not their economic values for the good. Two other studies go further by identifying the ethics behind this moral satisfaction. Stevens, Echeverria, Glass, Hager, and More (1991) asked respondents to interpret the meanings behind their willingness to pay for wildlife survival. They found that most respondents were expressing beliefs that wildlife have a right to exist independent of any benefit or harm to people. Baron and Spranca (1997) similarly tested the method and described environmental values as "protected values" arising from deontological ethics, not utilitarianism. Their interpretation also explained why the results of contingent valuation are sometimes not reliable:

> "Protected values are those that resist trade-offs with other values, particularly economic values. We propose that such values arise from deontological rules concerning action. People are concerned about their participation in transactions rather than just with the consequences that result. This proposal implies that protected values, defined as those that display trade-off resistance, will also tend to display quantity insensitivity, agent relativity, and moral obligation. People will also tend to experience anger at the thought of making trade-offs, and to engage in denial of the need for trade-offs through wishful thinking." (p. 1)

In light of these interpretations, it seems clear that the values of environmental valuation researchers are in conflict with those of the people they study.

Convergent Validity

While the construct validity of environmental valuations have been tested by others, this analysis focused on the convergent validity among the three methods. This refers to the degree to which different methods of valuation result in similar measures. It is, in essence, a measure of reliability among the different methods. The convergence of results is a necessary condition of validity, but not a sufficient one. That is, when convergence is found between methods, we still cannot claim the validity of the methods. We can only say that they are reliable. The meaning of the convergent results (i.e., the construct validity) is still open to debate. When, on the other hand, divergence is found between methods, then validity is jeopardized. The test does not say *which* methods are invalid, only that among the methods exhibiting divergence there is a validity problem.

Again, the results presented in Chapter 5 are mixed. Table 5.5 shows no significant difference between the average effect sizes from the travel cost and hedonic pricing methods. Their results converge[4], establishing a level of reliability among them. However, the results from contingent valuations are significantly different than each of the other two methods. Put together, these two results strongly suggest validity problems with contingent valuations[5], or at least confirm that contingent valuation measures a

[4] The small number of cases using the travel cost method makes this conclusion tenuous.
[5] However, the results do not *prove* that contingent valuation results are invalid. To answer which methods' results are invalid would require comparisons with actual market

different type of demand than the other two methods. Literature that evaluates the construct validity of contingent valuation are generally critical. Kealy, Dovidio, and Rockel (1988) compared contingent valuation results for familiar, market goods against actual market behavior for the goods and found low accuracy for the method. They concluded that the accuracy of results for less well defined, less familiar, environmental goods can be expected to be even lower. Johnston and Swallow (1999) found that the preferences expressed in their contingent valuation study conflict with the economic theory from which environmental valuation grew, casting doubt on the validity of the method. Others have sought to explain the causes of such divergence. Gregory, Lichtenstein, and Slovic (1993) evaluated contingent valuation from the perspective of behavioral decision research on human preferences and concluded that that the method imposes unrealistic cognitive demands on respondents, jeopardizing the validity of the method. Still, there is the possibility that the divergence between the methods' average effect sizes confirm Smith and Pattanayak's (2002) explanation for why there should be a divergence. Either way, at least one policy implication of this research is clear: these methods cannot be used interchangeably. Policies prescribing environmental valuation in their policy analyses must make distinctions between the methods of valuation and their respective results.

Defensible Uses

Given these mixed results, what uses of these valuation methods remain defensible? Others have described guidelines for the proper application of environmental

---

values, or some other measure of construct validity, which were beyond the scope of this research.

valuation tools, and the results from this study provide greater detail to these guidelines. Navrud and Pruckner (1997) developed a hierarchy of need for accuracy[6] in environmental valuation. As illustrated in Figure 6.2, accuracy of results is not needed when the purpose is simply to stimulate awareness and promote public discussion of the environmental changes. For example, Costanza et al. (1997) boldly estimated the economic value of the entire earth's ecosystem services, based upon benefit transfers of over 100 valuation studies, many of them contingent valuations. While the authors explicitly conceded the limitations of their estimate – including the biases in the valuation methods used in the original studies – they nonetheless presented their results to create awareness of the economic and environmental trade-offs humanity is making, consciously or not. Critics of their effort focus on the accuracy of the results, but admit that the effort brings attention to benefits provided by environmental services (Norgaard, Bode, & Values Reading Group, 1998). It showed, by order of magnitude, not precision, that the environment is worth far more than all the earth's combined gross domestic product. More accuracy is needed when evaluating projects and regulations, and when conducting environmental accounting, such as for cost-benefit analyses of proposed actions. The most accuracy is demanded in natural resource damage assessments, such as the case of valuing the environmental losses due to the *Exxon Valdez* oil spill for purposes of litigation and cost recovery.

The problem with these guidelines is that the accuracy of environmental valuations can rarely be directly confirmed because environmental goods are usually not traded in the market. Instead, prescriptions of valuation methods to these categories of

[6] In their study, the authors made no distinction between accuracy and validity and used both terms to refer to the degree to which valuation results match actual values of environmental goods. The same is done here.

Uses of environmental valuation	Accuracy demanded
▪ Natural resource damage assessment	high
▪ Environmental costing and accounting	↑
▪ Project evaluation; policy analysis; regulatory review	↑
▪ Stimulate awareness and public discussion	low

Figure 6.2: Demand for Accuracy in Environmental Valuation (adapted from Navrud & Pruckner, 1997)

uses can only be based upon indirect assessments of the methods' validity, such as the assessment of convergent validity in this research. Indeed, results of this study suggest that contingent valuation is best left for purposes of stimulating awareness and public discussion. The divergence of its average effect sizes from those of the other methods, combined with others' evidence of its construct validity problems, limit its defensible use to cases demanding only order of magnitude estimates of economic value. Its use in cost-benefit analyses, and other contexts demanding greater accuracy, is misleading at best (Diamond & Housman, 1994). Does this leave the hedonic pricing and travel cost methods to fulfill the other uses of environmental valuation? The evidence in this research does not invalidate these methods' results, but it also does not validate them because their results were only compared against each other, not against some objective benchmarks of values. Research is needed to test the construct validity of measuring economic values of bundled goods through these methods. If such research confirms the validity of the methods, then it is a much safer leap to assume that their use to value nonmarket environmental goods is also valid. Even so, they would still only measure direct use values, just one component of total economic value.

Norton (1995) also developed guidelines for economic valuation of environmental impacts. His environmental risk decision square conceptually maps environmental impacts against scales of space and time, and describes conditions under which economic decision rules are appropriate and inappropriate. In the square, an action with spatially wide and irreversible impacts affects the entire ecosystem beyond the length of a human life span. This is shown in Figure 6.3 as the "red zone," and such actions should not be valued by methods of individual preferences because they commit future generations to broad environmental impacts. At the other extreme, actions with local and short term impacts can be taken without affecting future generations or wide areas, so economic valuation methods can be appropriate.

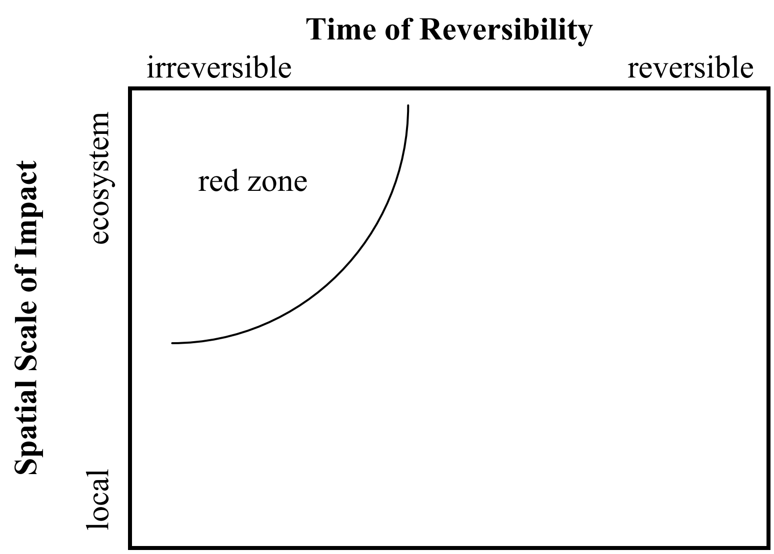


Figure 6.3: Environmental Risk Decision Square
(adapted from Norton, 1995)

The conclusions of this research accentuate Norton's prescriptions: the possible validity problems of the valuation methods further relegate their uses to low risk situations. Considering Navrud and Pruckner's and Norton's guidelines together, along with evidence from this research, the defensible uses of the environmental valuation methods are quite narrow. They can be appropriate in cases of local reversible impacts, and in such cases, hedonic pricing and travel cost methods are more reliable choices when accuracy is demanded. But, again, these two methods only capture direct use values and ignore all nonuse values. When accuracy is not needed, then estimates of total economic value from contingent valuation may stimulate discussions for local, reversible impacts. But when *both* accuracy and total economic value are needed, none of these three methods will suffice.

The results of this research also speak to the defensible uses of existing valuation studies in benefit transfers. The literature already identifies key study qualities to consider when transferring results from one context to another. They include the environmental good, the change in the good, the time in which the values were measured, the location of the study, the quality of the study, the demographics of the beneficiaries and the population studied, and the hypothetical markets in contingent valuations (Atkinson, Crocker, & Shogren, 1992; Brouwer, 2000; Smith, 1993; Whitehead & Hoban, 1999). All of these factors are presumed to moderate value, and the analyses presented in Chapters 4 and 5 help to qualify and quantify the effects of some of them.

The analysis confirms that different environmental goods do indeed moderate values. Air is the most transferable good, because of its homogeneity and the large number of studies valuing it. Homogeneity of the good reduces variance among the valuations, while the large number of studies increases the likelihood of finding existing studies that are comparable to the target situation in multiple ways. Animals and land, on the other hand, are the least transferable goods, because of the heterogeneity of these goods and the relatively few numbers of studies valuing each category of these goods.

This research is less conclusive about the importance of considering location and time in benefit transfers. While the average effect size in the United States is over twice that of other countries, the relatively few numbers of studies conducted outside the United States, and the diversity of countries represented, casts doubt on generalizability of this result. Within the United States, effect sizes were shown to be relatively stable across the states analyzed, except for California. Californians appear to place greater value on land improvements than the rest of the country, and less value on air improvements. Regarding the influence of the time of study, the evidence in this study shows it has relatively little effect on the measured values. While the preferences can reasonably be expected to change over time, the data in this study shows remarkably stable measures over the last three decades.

Most importantly, this research adds the method of valuation to the list of factors to consider when transferring benefits. Overall, contingent valuations produced effect sizes that averaged 40% to 55% less than those of the hedonic pricing and travel cost methods. However, the effect varies substantially by environmental good.

The evidence in this research also adds to the existing demands for better reporting in original valuation studies, so their results can be better judged in benefit transfers and meta-analyses (Carson, Flores, Martin, & Wright, 1996; Smith, 1993). In this research, 614 valuations of environmental changes were counted, but 30% of them were not usable in the analyses because they did not report sufficient information in their results. Commonly missing, fundamental information includes samples sizes and measures of variance. Other information that was often missing but could be quite helpful for secondary analyses include demographic characteristics of the samples such as income and age statistics, and study artifacts such as sampling methods and response rates.

Ultimately, benefit transfers are vulnerable to the same validity issues as the original studies, but to a greater degree because of the additional issue of the comparability of the circumstances in the original studies and the target situation. Thus benefit transfers should be limited to situations demanding low accuracy and relative economic values (Navrud & Pruckner, 1997).

Future Research with this Data

Further refinement of the defensible uses of these methods of environmental valuation is possible with additional analysis of this data set. Three broad issues appear particularly relevant, promising, and possible. One issue for deeper analysis is the moderating effects of valuation method. The current research confirmed and measured the influence of method, but it hides all the variance found within each method. Contingent valuation studies vary by the methods of elicitation, payment vehicles, amounts of information provided, components of value explicitly measured, response rates, etc., and many of these factors have been hypothesized or shown to affect results, as discussed in Chapter 5. A meta-analysis of the these variables could quantify these moderating effects, and explain more of the variance in contingent valuation results.

Similarly, hedonic pricing studies vary by functional form of the variables, the numbers and types of explanatory variables, the goods purchased, the source of dependent variable, etc., and the data set could test the effects of these factors on hedonic pricing results. Also, with more sensitive coding of the magnitudes of environmental change, including temporal and spatial boundaries, a "meta-scope test" could be performed to test the methods' sensitivity to the scope of goods (Gregory, Mendelsohn, & Moore, 1989).

Another issue for refinement is the moderating effects of environmental goods. Further meta-analysis of it could reveal preferences people hold for variants of goods. For example, while others have attempted to measure the values of surface water in different forms (Mahan, Polasky, & Adams, 2000), a meta-analysis of this data set could measure the relative preferences people hold for ground water versus surface water. Such information could be important to water resources planning and development. Similarly, the relative preferences of different forms of land could be assessed with this data set to inform land use planning and development.

This data set could also be analyzed to compare use values and nonuse values measured by contingent valuation. The adaptability of the method to the different components of economic value is its strength to proponents, but may also explain its problems with reliability (Diamond & Housman, 1994; More, Averill, & Stevens, 1996; Rosenthal, Nelson, & Kopp, 1992). Because it can measure different components of economic value, its results can be expected to exhibit greater variance, thereby reducing its average effect size. On the other hand, average effect sizes from contingent valuations can also be expected to be larger, because they are presumed to measure more components of value. A meta-analysis of the components of economic value might sort use values from nonuse values and explain much of the variance in results.

What They Don't Tell Us, But Should

What environmental valuations aim to tell us is a limited, economic view of value. There are two major problems with this. First, as described above, the methods do not do a great job of doing so. Contingent valuation is fraught with validity problems, while the hedonic pricing and travel cost methods measure just a slice of economic value. More importantly, economic values are not as relevant to the environment as the popularity of environmental valuations suggests. Subjects in environmental valuation studies make decisions based upon the advancement of their personal well being, but the environment is not a commodity to be traded for the purpose of raising personal welfare (Jakobsson & Dragun, 2001; Kassiola, 2003b; Rees, 1998). Instead, the environment elicits many other dimensions of value that have been vying for a substantial voice in public policy. Some of them have had notable success in being codified[7], but none has displaced economic value in popularity, or approached the well defined, quantitative methods of measurement enjoyed by economic values.

Even proponents of these methods are quick to note that other dimensions of value are relevant to environmental goods and ought to be considered in policy decisions (Hanley, 1992). But oftentimes, the inclusion of economic measures of value actually causes the loss of other measures because it reduces and limits the discussion by presenting what is presumed to be compact, commensurate information (Vatn &

[7] For example, environmental justice is codified in Executive Order 12989, and public opinions on environmental impacts in 40CFR1502.

Bromley, 1994). To the extent that environmental valuations limit the discussions to only economic considerations, they are detrimental to environmental decision making. Instead, valuations of the environment ought to facilitate expressions of many dimensions of value (Norgaard, 1985; Paehlke, 2003). The literature again provides guidance on other values relevant to environmental decision making. Interestingly, it lists values that are in tension with foundations of economic theory:

- Communal, public values (not just individuals' private values). Stone (1997) strongly argues that public policy decisions, such as those affecting the environment, are inherently communal and are misdirected when a purely economic perspective is taken. In such a framework, the individual is sovereign, and each person narrowly seeks to advance personal interests. Bozeman (2002) claims that such an approach leads to "public value failure," failures of efficient markets to produce outcomes reflecting public values. "What most economic approaches to public value have in common is that they are less a reflection of public value than the private value of things public" (p. 146). Instead, Stone argues that public policy decisions are best understood and directed when the community – or the "polis" as she calls it – is sovereign and individuals make decisions together with a focus on public interest. Several others agree (e.g., Gowdy, 1997; Norton, Costanza, & Bishop, 1998; Sagoff, 1997).

- Cooperative values (not just competitive values). Closely related to communal values are cooperative values. While the former identifies a unit of analysis (community), the latter identifies a process. Markets tend to generate competitive behavior and discourage cooperative behavior, but this is an outcome of market processes, not a reflection of human nature (McLaughlin, 2003). Thus, Shabman and Stephenson (1996) claim that any valuation method's validity and usefulness should be judged by its ability to facilitate cooperative decision making. They argue that values for environmental goods are constructed in dialogue and interactions, not held individually in some a priori preference logic.

- Deontological values (not just utilitarian values). Milbrath (2003) argues that economic valuation abdicates moral choices on the environment to markets, when justice is a more appropriate value to measure. Sagoff (1981) agrees. He states that judgments on environmental choices ought to reflect deontological values, not just utilitarian ones. Kempton, Boster, and Hartley (1995) provide empirical evidence supporting this. In their public opinion survey on environmental values, they found that only a small minority of Americans take a purely utilitarian view of environmental goods, even among people in the timber industry who directly benefit from the commodity value of trees. Environmental justice is an example of deontological values considered in environmental decision making, but its lack of a clear operational measure – as well as the controversial politics surrounding it – has kept it from a position of prominence in environmental decision making.

- Environmental quality (not just personal wealth). When personal gain is the measure of value for common-pool[8] environmental goods, the resources' degradation is eminent (Hardin 1968). To counter this tendency, environmental quality must be included in the measures of value. For example, the value of trees may be monetarily quantified by their potential market value as paper or lumber. Such analysis, however, would neglect their values in preventing soil-erosion, using carbon dioxide and generating oxygen, and providing habitat for wildlife (Brower, 1995). Measuring these latter values identify an environmental objective of decision making that complements the welfare objective of the economic view. Gordon Brady, then a member of the Council on Environmental Quality under President Ronald Reagan, understood these things when his boss called for cost-benefit decision rules for government actions, in Executive Order 12291. To balance the purely economic view, Brady (1983) prescribed direct measures of environmental quality to be included in calculations of environmental values.

Taken together, the relevant measures of environmental values include economic terms, non-economic social terms, and environmental terms. They represent a pluralistic approach to replace one that has been dominated by, if not limited to, economic considerations (Gowdy, 1997; Norgaard, 1985).

[8] Common-pool resources are those that are non-excludable in consumption, yet exhaustible.

Alternative Methods

These prescriptions for what environmental valuations ought to tell us begs the question, "How can these components of value be measured?" Finding a single measure to capture all aspects of relevant values is unlikely and perhaps undesirable, but options are available to capture the relevant values with a combination of measures. Alternative methods of value measurement can be grouped by two variables: component of value and unit of analysis. Several methods are summarized in Table 6.1 and select ones are introduced below.

Many methods have been developed to measure the economic component of value. Besides those evaluated in this study, others observe groups and environmental qualities. Multi-attribute value elicitation is a modification of contingent valuation that is explicit in capturing several aspects of total economic value (McDaniels & Roessler, 1998). Doing so makes the interpretation of the results less problematic, but the method is vulnerable to many of the same problems as contingent valuation. Deliberative contingent valuation acknowledges the problems of applying utilitarian ethics to environmental decision making, but does not abandon economic valuation. Instead, it assumes that contingent valuation can facilitate collective judgments of value by making the process discursive (Sagoff, 1998). The jury-like approach is still aimed at economic valuation of environmental goods, but the task is couched as a social preference, not individual. Two other methods are creative in their analysis of environmental measures in economic terms. The environmental annuities method (Unsworth & Bishop, 1994) and the net environmental benefit analysis (NEBA) (J. Fittipaldi, personal communication, 2001) essentially apply methods of discounted cash flow analysis and cost-benefit

Table 6.1: Methods of Valuation by Component of Value and Unit of Analysis

Unit of Analysis	Component of Value: economic terms	non-economic social terms	environmental terms
individuals	▪ travel cost method ▪ hedonic pricing ▪ contingent valuation ▪ multi-attribute value elicitation	▪ hearing ▪ social survey ▪ decision pathway ▪ narrative valuation	▪ environmental testimony
groups	▪ deliberative contingent valuation	▪ focus group ▪ public debate ▪ discursive ethics forum ▪ citizens' jury ▪ environmental justice	▪ environmental panel
environment	▪ environmental annuities ▪ NEBA ▪ multiple attribute decision making	▪ environmental justice	▪ index of biologic integrity ▪ indicators ▪ NEBA
other	▪ multiple attribute decision making	▪ content analysis ▪ historiography	

analysis to environmental changes, but they do so in ecological units of measure. Thus, they avoid all the controversy over monetary conversion of environmental changes. NEBA was developed by the U.S. Army to prioritize remediation efforts and has been endorsed by the U.S. Environmental Protection Agency and the U.S. Department of the Interior ("Army…," 2001). Multiple attribute decision making is a framework for evaluating environmental alternatives based on operations research (Prato, 1999). Like monetary environmental valuation, its criterion of evaluation is efficiency, but unlike monetary valuation, it considers multiple dimensions of value.

Government bodies and social sciences have developed popular methods of measuring social values and facilitating social decision making, such as hearings, social surveys, focus groups, and debates. Added to this list are several methods of social value measurement developed specifically for environmental decision making. Decision pathways is a surveying technique that not only measures preferences, but also facilitates the construction of expressed values through series of linked questions (Gregory et al., 1997). Participants in the method are asked a sequence of questions that is determined by their answers to previous questions. Several different question-and-answer pathways are possible, each one representing different values underlying their stated preferences. Narrative valuation is a method designed to help respondents consider different values in their development of environmental judgments (Satterfield, Slovic, & Gregory, 2000). As its name suggests, it does so through qualitative narratives of the problem. Discursive ethics, a conceptual framework stemming from the theoretical roots of the Frankfurt School of critical theory, deconstructs environmental value claims from multiple disciplines to identify and compare underlying value distinctions (O'Hara, 1996). It manifests itself as open discourse among stakeholders which can take the forms of conflict resolution, mediation, etc. A few methods of social value measurement are noteworthy for their creativity or unique perspective. Citizens' jury provides an interesting environmental decision making method loosely based on the idea of a court jury (Aldred & Jacobs, 2000). Representative jurors are selected from the public to hear arguments for environmental alternatives. Evidence is presented by experts on the alternatives, the jury deliberates, and their decision is expected to carry some authority. Content analysis is based on an established claim that news media coverage of issues are correlated with public values. Applied to environmental valuation, the method examines topics and prevalence of environmental issues in the news media to assess values for them (Bengston, Fan, & Celarier, 1999). Finally, historiography recognizes that the history of an environmental good is a component of its value (Goodin, 2003). Recording of its history preserves that value.

The environmental sciences are responsible for developing the direct measures of environmental quality, such as environmental indicators and indices. Environmental indicators are physical measures of select environmental goods or conditions that are meant to represent broader environmental quality (Nyborg, 2000). For example, the prevalence of a key species may be an indicator of the health of its ecosystem. A more direct approach recognizes biologic integrity as a measure of ecosystem health (Karr, 2002), and indices of biologic integrity have been developed for different ecosystems (e.g., Karr & Rossano, 2001; Lyons, Navarro-Perez, Cochran, Santana, & Guzman-Arroyo, 1995). Components of a specific riparian index, for example, might include counts of kinds of organisms at a site (biodiversity), relative abundance of predators, fish assemblages, etc. Such indices provide a continuum of values beyond the dichotomous categories of "impaired" and "unimpaired" that are often used to inform decisions.

Maguire (2002) claims that good methods of measuring environmental values are those that facilitate public decision making by 1) articulating those values, 2) making tradeoffs among conflicting goals, 3) understanding sources of conflict, 4) solving distributional problems, 5) integrating values with technical analysis, and 6) anticipating future consequences. By themselves, the three methods of economic valuation in this research do ostensibly well in facilitating tradeoffs, but they fail miserably on all other counts. However, a pluralistic approach to environmental valuation that incorporates some of the alternative methods introduced here can meet all of these criteria, and thus improve environmental policy decision making processes.

The Lesson from the *Exxon Valdez*

The conclusions of this research have been anticipated by policy practitioners in the last decade. Let us return to the case of the *Exxon Valdez* to observe the rise and decline of the practice. The ship spilled its oil cargo in Prince William Sound in 1989, and in the ensuing battle over cost recovery, private and government organizations conducted environmental valuation studies to estimate the economic value of the environmental damages. The case brought much attention to the practice of environmental valuation, and eventually a government endorsement of contingent valuation in 1993. But it also increased the existing controversies surrounding the methods, partly because the estimated values of the environmental damages varied by billions of dollars (House of Representatives Committee on Merchant Marine Fisheries, 1991), and partly because some observers were perplexed by the measurement of environmental damages in monetary terms (Keeble, 1999). Instead of clarifying the magnitude of the damages, many observers felt environmental valuation had made it more obscure. Under such controversy, the National Oceanic and Atmospheric Administration ceased its practice of environmental valuation in 1995, just two years after it had endorsed it (Rogers, 1995). NOAA announced it would no longer monetize the values of dead animals or damaged ecosystems following oil spills. Instead, it would measure environmental impacts in ecological terms and engage the public to determine how to restore the ecosystem. This restoration-based approach acknowledges the vulnerabilities and limitations of monetary valuation methods (especially for natural resource damage assessment), and replaces it with methods that are more community oriented, cooperative, and focused on environmental quality. The results of this research provides econometric evidence to this experiential lesson of the *Exxon Valdez* case. They confirm that methods of monetizing environmental values should have a much smaller role in environmental decision making than once thought and prescribed by policy practitioners.

APPENDIX A

CODE SHEET

META-ANALYSIS OF ENVIRONMENTAL VALUATION – CODE SHEET

Case number:
- Author(s):
- Year of publication:
- Title:
- Journal:
- V(N): pp.:

* * * * * * * *

Unit of analysis number:

Subjects
- Data year(s):
- Locations of data sources:
- Local, regional, national, international population:
- Original sample size:
- Response rate:
- Usable responses (n) / subsets (n1, n2):
- Sampling method (probability/non-probability):
- Mean income / unit:

Change in environmental good / service
- Good / service:
- Damage assessment?
- Baseline quality/level:
- Quantitative or qualitative measure of change:
- Description of change:
- Incremental or holistic change:
- Number and types of measures used:
- Temporal boundary of change:
- Spatial boundary of change (local, regional, national, international):
- Type of effect (human, ecological, etc.):

Valuation

<table>
<tr><th>☐ Travel cost</th><th>☐ Hedonic pricing</th><th>☐ Contingent Valuation</th></tr>
<tr>
<td valign="top">

Individual/zonal:
DV:
visits
visit or not
other:
Medium:
Phone
Mail
In-person
Focus group
Other:

</td>
<td valign="top">

Weakly complementary good:
Residential
Agricultural
Commercial
Other:
Source of DV:
Sales
Census tract data
Appraisals
Other:
DV:
Functional form:
DV:
Env IV:
Number of IV:
Site:
Neighborhood:
Environmental:
Other:

</td>
<td valign="top">

Elicitation method:
Open-ended / direct
Discrete choice
Referendum w/o follow-up
Referendum w/ follow-up
Iterative bidding
Other:
Payment vehicle:
Taxes
Fees
Other:
DV (WTP, WTA):
Medium:
Phone
Mail
In-person
Focus group
Other:
Separation of estimates:
Hypothetical market:
Description included?
words

</td>
</tr>
</table>

- Value, with descriptive statistics:
- Effect size derivation:
- Effect size, d:
- Comparison with other methods?

Conclusions/messages of the article:

APPENDIX B

STUDIES INCLUDED IN THIS META-ANALYSIS

Adamowicz, W., Boxall, P., Williams, M., & Louviere, J. (1998). Stated preference approaches for measuring passive use values: choice experiments and contingent valuation. *American Journal of Agricultural Economics, 80*(1), 64-75.

Adams, R.M., Bergland, O., Musser, W.N., Johnson, S.L., & Musser, L.M. (1989). User fees and equity issues in public hunting expenditures: the case of ring-necked pheasant in Oregon. *Land Economics, 65*(4), 376-385.

Alberini, A., Kanninen, B., & Carson, R.T. (1997). Modeling response incentive effects in dichotomous choice contingent valuation data. *Land Economics, 73*(3), 309-324.

An, M.Y. (2000). A semiparametric distribution for willingness to pay and statistical inference with dichotomous choice contingent valuation data. *American Journal of Agricultural Economics, 82*(3), 487-500.

Anex, R.P. (1995). A travel-cost method of evaluating household hazardous waste disposal services. *Journal of Environmental Management, 45*, 189-198.

Arimah, B.C. (1996). Willingness to pay for improved environmental sanitation in a Nigerian City. *Journal of Environmental Management, 48*, 127-138.

Barnard, J.R. (1978). Externalities from urban growth: the case of increased storm runoff and flooding. *Land Economics, 54*(3), 298-315.

Beach, E.D. & Carlson, G.A. (1993). A hedonic analysis of herbicides: do user safety and water quality matter? *American Journal of Agricultural Economics, 75*(3), 612-623.

Bergstrom, J.C., Stoll, J.R., & Randall, A. (1990). The impact of information on environmental commodity valuation decisions. *American Journal of Agricultural Economics, 72*(3), 614-621.

Beron, K.J., Murdoch, J.C., & Thayer, M.A. (1999). Hierarchical linear models with application to air pollution in the south coast air basin. *American Journal of Agricultural Economics, 81*(5), 1123-1127.

Berrens, R.P., Bohara, A.K., Silva, C.L., Brookshire, D., & McKee, M. (2000). Contingent values for New Mexico instream flows: with tests of scope, group-size reminder and temporal reliability. *Journal of Environmental Management, 58*, 73-90.

Berry, B.J.L. & Bednarz, R.S. (1975). A hedonic model of prices and assessments for single-family homes: does the assessor follow the market or the market follow the assessor? *Land Economics, 51*, 21-40.

Bjornstad, D., Cummings, R., & Osborne, L. (1997). A learning design for reducing hypothetical bias in the contingent valuation method. *Environmental and Resource Economics, 10*, 207-221.

Bockstael, N.E., Hanemann, W.M., & Kling, C.L. (1987). Estimating the value of water quality improvements in a recreational demand framework. *Water Resources Research, 23*(5), 951-960.

Bolitzer, B. & Netusil, N.R. (2000). The impact of open spaces on property values in Portland, Oregon. *Journal of Environmental Management, 59*, 185-193.

Boman, M., Bostedt, G., & Kristrom, B. (1999). Obtaining welfare bounds in discrete-response valuation studies: a non-parametric approach. *Land Economics, 75*(2), 284-294.

Bonnieux, F. & LeGoffe, P. (1997). Valuing benefits of landscape restoration: a case study of the Cotentin in Lower-Normandy, France. *Journal of Environmental Management, 50*, 321-333.

Bouwes, N.W. & Schneider, R. (1979). Procedures in estimating benefits of water quality change. *American Journal of Agricultural Economics, 61*(3), 535-539.

Bowker, J.M. & Stoll, J.R. (1988). Use of dichotomous choice nonmarket methods to value the whooping crane resource. *American Journal of Agricultural Economics, 70*(2), 372-381.

Boyle, K.J. (1989). Commodity specification and the framing of contingent-valuation questions. *Land Economics, 65*(1), 57-63.

Boyle, K.J. & Bishop, R.C. (1987). Valuing wildlife in benefit-cost analysis: a case study involving endangered species. *Water Resources Research, 23*(5), 943-950.

Boyle, K.J. & Bishop, R.C. (1988). Welfare measurements using contingent valuation: a comparison of techniques. *American Journal of Agricultural Economics, 70*(1), 20-28.

Boyle, K.J., Bishop, R.C., & Welsh, M.P. (1985). Starting point bias in contingent valuation bidding games. *Land Economics, 61*(2), 188-194.

Boyle, K.J., Johnson, F.R., McCollum, D.W., Desvousges, W.H., Dunford, R.W., & Hudson, S.P. (1996). Valuing public goods: discrete versus continuous contingent-valuation responses. *Land Economics, 72*(3), 381-396.

Boyle, K.J., Poor, P.J., & Taylor, L.O. (1999). Estimating the demand for protecting freshwater lakes from eutrophication. *American Journal of Agricultural Economics, 81*(5), 1118-1122.

Brookshire, D.S., Eubanks, L.S., & Randall, A. (1983). Estimating option prices and existence values for wildlife resources. *Land Economics, 59*(1), 1-15.

Brown, R.J. (1976). A study of the impact of the Wetlands Easement Program on agricultural land values. *Land Economics, 52*(4), 509-517.

Brown, T.C., Barro, S.C., Manfredo, M.J., & Peterson, G.L. (1995). Does better information about the good avoid the embedding effect? *Journal of Environmental Management, 44*, 1-10.

Brown, T.C., Champ, P.A., Bishop, R.C., & McCollum, D.W. (1996). Which response format reveals the truth about donations to a public good? *Land Economics 72*(2), 152-166.

Brown, T.C. & Duffiend, J.W. (1995). Testing part-whole valuation effects in contingent valuation of instream flow protection. *Water Resources Research, 31*(9), 2341-2351.

Brucato, P.F., Murdoch, J.C., & Thayer, M.A. (1990). Urban air quality improvements: a comparison of aggregate health and welfare benefits to hedonic price differentials. *Journal of Environmental Management, 30*, 265-279.

Carson, R.T., Hanemann, W.M., Kopp, R.J., Krosnick, J.A., Mitchell, R.C., Presser, S. Ruud, P.A., Smith, V.K., Conaway, M., & Martin, K. (1997). Temporal reliability of estimates from contingent valuation. *Land Economics, 73*(2), 151-163.

Caulkins, P.P., Bishop, R.C., & Bouwes, N.W. (1986). The travel cost model for lake recreation: a comparison of two methods for incorporating site quality and substitution effects. *American Journal of Agricultural Economics, 68*(2), 291-297.

Champ, P.A. & Bishop, R.C. (2001). Donation payment mechanisms and contingent valuation: an empirical study of hypothetical bias. *Environmental and Resource Economics, 19*, 383-402.

Chattopadhyay, S. (1999). Estimating the demand for air quality: new evidence based on the Chicago housing market. *Land Economics, 75*(1), 22-38.

Choe, K.C., Whittington, D., & Lauria, D.T. (1996). The economic benefits of surface water quality improvements in developing countries: a case study of Davao, Philippines. *Land Economics, 72*(4), 519-537.

Christie, M. (2001). A comparison of alternative contingent valuation elicitation treatments for the evaluation of complex environmental policy. *Journal of Environmental Management, 62*, 255-269.

Clayton, C. & Mendelsohn, R. (1993). The value of watchable wildlife: a case study of McNeil River. *Journal of Environmental Management, 39*, 101-106.

Cooper, J.C. (2000). Nonparametric and semi-nonparametric recreational demand analysis. *American Journal of Agricultural Economics, 82*(2), 451-462.

Cooper, J. & Loomis, J. (1993). Testing whether waterfowl hunting benefits increase with greater water deliveries to wetlands. *Environmental and Resource Economics, 3*, 545-561.

Cordell, H.K. & Bergstrom, J.C. (1993). Comparison of recreation use values among alternative reservoir water level management scenarios. *Water Resources Research, 29*(2), 247-258.

Correll, M.R., Lillydahl, J.H., & Singell, L.D. (1978). The effects of greenbelts on residential property values: some findings on the political economy of open space. *Land Economics, 54*(2), 207-217.

Crocker, T.D. (1985). On the value of the condition of a forest stock. *Land Economics, 61*(3), 244-254.

Croke, K., Fabian, R., & Brenniman, G. (1986). Estimating the value of natural open space preservation in an urban area. *Journal of Environmental Management, 23*, 317-324.

Crouter, J.P. (1987). Hedonic estimation applied to a water rights market. *Land Economics, 63*(3), 259-271.

Cummings, R.G., Ganderton, P.T., & McGuckin, T. (1994). Substitution effects in CVM values. *American Journal of Agricultural Economics, 76*(2), 205-214.

Dalecki, M.G., Whitehead, J.C., & Blomquist, G.C. (1993). Sample non-response bias and aggregate benefits in contingent valuation: and examination of early, late and non-respondents. *Journal of Environmental Management, 38*, 133-143.

Darling, A.H. (1973). Measuring benefits generated by urban water parks. *Land Economics, 49*, 22-34.

Daubert, J.T. & Young, R.A. (1981). Recreation demands for maintaining instream flows: a contingent valuation approach. *American Journal of Agricultural Economics, 63*(4), 666-676.

Diamond, D.B. (1980). The relationship between amenities and urban land prices. *Land Economics, 56*(1), 21-32.

Duffield, J.W., Neher, C.J., & Brown, T.C. (1992). Recreation benefits of instream flow: application to Montana's Big Hole and Bitterroot Rivers. *Water Resources Research, 28*(9), 2169-2181.

duVair, P. & Loomis, J. (1993). Household's valuation of alternative levels of hazardous waste risk reductions: an application of the referendum format contingent valuation method. *Journal of Environmental Management, 39*, 143-155.

Earnhart, D. (2001). Combining revealed and stated preference methods to value environmental amenities at residential locations. *Land Economics, 77*(1), 12-29.

Edmonds Jr., R.G. (1985). Some evidence on the intertemporal stability of hedonic price functions. *Land Economics, 61*(4), 445-451.

Edwards, S.F. & Anderson, G.D. (1987). Overlooked biases in contingent valuation studies: some considerations. *Land Economics, 63*(2), 168-178.

Eiswerth, M.E., Englin, J., Fadali, E., & Shaw, W.D. (2000). The value of water levels in water-based recreation: a pooled revealed preference/contingent behavior model. *Water Resources Research, 36*(4), 1079-1086.

Ekstand, E.R. & Loomis, J. (1998). Incorporating respondent uncertainty when estimating willingness to pay for protecting critical habitat for threatened and endangered fish. *Water Resources Research, 34*(11), 3149-3155.

Englin, J. & Lambert, D. (1995). Measuring angling quality in count data models of recreational fishing. *Environmental and Resource Economics, 6*, 389-399.

Epp, D.J. & Al-Ani, K.S. (1979). The effect of water quality on rural nonfarm residential property values. *American Journal of Agricultural Economics, 61*(3), 529-534.

Faux, J. & Perry, G.M. (1999). Estimating irrigation water value using hedonic price analysis: a case study in Malheur County, Oregon. *Land Economics, 75*(3), 440-452.

Folland, S.T. & Hough, R.R. (1991). Nuclear power plants and the value of agricultural land. *Land Economics, 67*(1), 30-36.

Garrod, G. & Willis, K. (1992). The amenity value of woodland in Great Britain: a comparison of economic estimates. *Environmental and Resource Economics, 2*, 415-434.

Garrod, G.D. & Willis, K.G. (1992). Valuing goods' characteristics: an application of the hedonic price method to environmental attributes. *Journal of Environmental Management, 34*, 59-76.

Garrod, G. & Willis, K. (1994). An economic estimate of the effect of a waterside location on property values. *Environmental and Resource Economics, 4*, 209-217.

Giraud, K.L., Loomis, J.B., & Johnson, R.L. (1999). Internal and external scope in willingness-to-pay estimates for threatened and endangered wildlife. *Journal of Environmental Management, 56*, 221-229.

Goldar, B. & Misra, S. (2001). Valuation of environmental goods: correcting for bias in contingent valuation studies based on willingness-to-accept. *American Journal of Agricultural Economics, 83*(1), 150-156.

Graves, P., Murdoch, J.C., Thayer, M.A., & Waldman, D (1988). The robustness of hedonic price estimation: urban air quality. *Land Economics, 64*(3), 220-233.

Gregory, R. (2000). Valuing environmental policy outcomes: a case study comparison of multiattribute and contingent valuation survey methods. *Land Economics, 76*(2), 151-173.

Gregory, R., Lichtenstein, S., Brown, T.C., Peterson, G.L., & Slovic, P. (1995). How precise are monetary representations of environmental improvements? *Land Economics, 71*(4), 462-473.

Haab, T.C. (1999). Nonparticipation or misspecification? *Environmental and Resource Economics, 14*, 443-461.

Haab, T.C. & McConnell, K.E. (1998). Referendum models and economic values: theoretical, intuitive, and practical bounds on willingness to pay. *Land Economics, 74*(2), 216-229.

Hailu, A., Adamowicz, W.L., & Boxall, P.C. (2000). Complements, substitutes, budget constraints and valuation. *Environmental and Resource Economics, 16*, 51-68.

Halvorsen, B. (1996). Ordering effects in contingent valuation surveys. *Environmental and Resource Economics, 8*, 485-499.

Halvorsen, B. (2000). Comparing ranking and contingent valuation for valuing human lives, applying nested and non-nested logit models. *Environmental and Resource Economics, 17*, 1-19.

Halvorsen, B. & Saelensminde, K. (1998). Differences between willingness-to-pay estimates from open-ended and discrete-choice contingent valuation methods: the effects of heteroscedasticity. *Land Economics, 74*(2), 262-282.

Hamilton, S.W. & Schwann, G.M. (1995). Do high voltage electric transmission lines affect property values? *Land Economics, 71*(4), 436-444.

Hanemann, M., Loomis, J., & Kanninen, B. (1991). Statistical efficiency of double-bounded dichotomous choice contingent valuation. *American Journal of Agricultural Economics, 73*(4), 1255-1263.

Hanley, N., Spash, C. & Walker, L. (1995). Problems in valuing the benefits of biodiversity protection. *Environmental and Resource Economics, 5*, 249-272.

Harris, B.S. (1984). Contingent valuation of water pollution control. *Journal of Environmental Management, 19*, 199-208.

Hoehn, J.P. (1991). Valuing the multidimensional impacts of environmental policy: theory and methods. *American Journal of Agricultural Economics, 73*(2), 289-299.

Hoevenagel, R. (1996). The validity of the contingent valuation method: perfect and regular embedding. *Environmental and Resource Economics, 7*, 57-78.

Howe, C.W. & Smith, M.G. (1993). Incorporating public preferences in planning urban water supply reliability. *Water Resources Research, 29*(10), 3363-3369.

Huhtala, A. (1999). How much do money, inconvenience and pollution matter? Analyzing households' demand for large-scale recycling and incineration. *Journal of Environmental Management, 55*, 27-38.

Huhtala, A. (2000). Binary choice valuation studies with heterogeneous preferences regarding the program being valued. *Environmental and Resource Economics, 16*, 263-279.

Irwin, E.G. & Bockstael, N.E. (2001). The problem of identifying land use spillovers: measuring the effects of open space on residential property values. *American Journal of Agricultural Economics, 83*(3), 698-704.

Jakobsson, K.M. & Dragun, A.K. (2001). The worth of a possum: valuing species with the contingent valuation method. *Environmental and Resource Economics, 19*, 211-227.

Johansson, P.O., Kristrom, B., & Mattsson, L. (1988). How is the willingness to pay for moose hunting affected by the stock of moose? An empirical study of moose-hunters in the county of Vasterbotten. *Journal of Environmental Management, 26*, 163-171.

Johnson, N.S. & Adams, R.M. (1988). Benefits of increased streamflow: the case of the John Day River steelhead fishery. *Water Resources Research, 24*(11), 1839-1846.

Johnston, R.J. & Swallow, S.K. (1999). Asymmetries in ordered strength of preference models: implications of focus shift for discrete-choice preference estimation. *Land Economics, 75*(2), 295-310.

Jordan, J.L. & Elnagheeb, A.H. (1993). Willingness to pay for improvements in drinking water quality. *Water Resources Research, 29*(2), 237-245.

Jorgensen, B.S., Syme, G.J., Bishop, B.J., & Nancarrow, B.E. (1999). Protest responses in contingent valuation. *Environmental and Resource Economics, 14*, 131-150.

Kanninen, B.J. & Khawaja, M.S. (1995). Measuring goodness of fit for the double-bounded logit model. *American Journal of Agricultural Economics, 77*(4), 885-890.

Kaoru, Y. (1993). Differentiating use and nonuse values for coastal pond water quality improvements. *Environmental and Resource Economics, 3*, 487-494.

Kealy, M.J. & Turner, R.W. (1993). A test of the equality of closed-ended and open-ended contingent valuations. *American Journal of Agricultural Economics, 75*(2), 321-331.

Kerr, G.N. & Cullen, R. (1995). Public preferences and efficient allocation of a Possum-control budget. *Journal of Environmental Management, 43*, 1-15.

Kiel, K.A. (1995). Measuring the impact of the discovery and cleaning of identified hazardous waste sites on house values. *Land Economics, 71*(4), 428-435.

King, D.A. & Sindemn, J.A. (1988). Influence of soil conservation on farm land values. *Land Economics, 64*(3), 242-255.

Kirshner, D. & Moore, D. (1989). The effect of San Francisco Bay water quality on adjacent property values. *Journal of Environmental Management, 27*, 263-274.

Kramer, R.A. & Mercer, D.E. (1997). Valuing a global environmental good: U.S. residents' willingness to pay to protect tropical rain forests. *Land Economics, 73*(2), 196-210.

Kriesell, W., Randall, A., & Lichtkoppler, F. (1993). Estimating the benefits of shore erosion protection in Ohio's Lake Erie housing market. *Water Resources Research, 29*(4), 795-801.

Kristrom, B. (1990). A non-parametric approach to the estimation of welfare measures in discrete response valuation studies. *Land Economics, 66*(2), 134-139.

Kristrom, B. (1993). Comparing continuous and discrete contingent valuation questions. *Environmental and Resource Economics, 3*, 63-71.

Kristrom, B. (1997). Spike models in contingent valuation. *American Journal of Agricultural Economics, 79*(3), 1013-1023.

Kwak, S. & Russell, C.S. (1994). Contingent valuation in Korean environmental planning: a pilot application to the protection of drinking water quality in Seoul. *Environmental and Resource Economics, 4*, 511-526.

Langford, I.H. (1994). Using a generalized linear mixed model to analyze dichotomous choice contingent valuation data. *Land Economics, 70*(4), 507-514.

Langford, I.H., Bateman, I.J., Jones, A.P., Langford, H.D., & Georgiou, S. (1998). Improved estimation of willingness to pay in dichotomous choice contingent valuation studies. *Land Economics, 74*(1), 65-75.

Langford, I.H., Bateman, I.J., & Langford, H.D. (1996). A multilevel modeling approach to triple-bounded dichotomous choice contingent valuation. *Environmental and Resource Economics, 7*, 197-211.

Langford, I.H., Kontogianni, A., Skourtos, M.S., Georgiou, S., & Bateman, I.J. (1998). Multivariate mixed models for open-ended contingent valuation data. *Environmental and Resource Economics, 12*, 443-456.

Laughland, A.S., Musser, W.N., Shortle, J.S., & Musser, L.M. (1996). Construct validity of averting cost measures of environmental benefits. *Land Economics, 72*(1), 100-112.

Lee, H.C. & Chun, H.S. (1999). Valuing environmental quality change on recreation hunting in Korea: a contingent valuation analysis. *Journal of Environmental Management, 57*, 11-20.

LeGoffe, P. (1995). The benefits of improvements in coastal water quality: a contingent approach. *Journal of Environmental Management, 45*, 305-317.

LeGoffe, P. (2000). Hedonic pricing of agriculture and forestry externalities. *Environmental and Resource Economics, 15*, 397-401.

Leon, C.J. (1996). Double bounded survival values for preserving the landscape of natural parks. *Journal of Environmental Management, 46*, 103-118.

Li, C. (1996). Semiparametric estimation of the binary choice model for contingent valuation. *Land Economics, 72*(4), 462-473.

Li, M.M. & Brown, H.J. (1980). Micro-neighborhood externalities and hedonic housing prices. *Land Economics, 56*(2), 125-141.

Lichtenberg, E. & Zimmerman, R. (1999). Farmers' willingness to pay for groundwater protection. *Water Resources Research, 35*(3), 833-841.

Livengood, K.R. (1983). Value of big game from markets for hunting leases: the hedonic approach. *Land Economics, 59*(3), 287-291.

Lockwood, M. (1999). Preference structures, property rights, and paired comparisons. *Environmental and Resource Economics, 13*, 107-122.

Lockwood, M., Loomis, J., & DeLacy, T. (1993). A contingent valuation survey and benefit-cost analysis of forest preservation in East Gippsland, Australia. *Journal of Environmental Management, 38*, 233-243.

Loehman, E.T., Park, S., & Boldt, D. (1994). Willingness to pay for gains and losses in visibility and health. *Land Economics, 70*(4), 478-498.

Loomis, J. (1986). Economic losses to recreational fisheries due to small-head hydro-power development: a case study of the Henry's Fork in Idaho. *Journal of Environmental Management, 22*, 85-94.

Loomis, J.B. (1987). Balancing public trust resources of Mono Lake and Los Angeles' water right: an economic approach. *Water Resources Research, 23*(8), 1449-1456.

Loomis, J.B. (1987). Expanding contingent value sample estimates to aggregate benefit estimates: current practices and proposed solutions. *Land Economics, 63*(4), 396-402.

Loomis, J.B. (1989). Estimation of and variation in site specific marginal values for recreational fisheries. *Journal of Environmental Management, 29*, 183-191.

Loomis, J.B. (1989). Test-retest reliability of the contingent valuation method: a comparison of general population and visitor responses. *American Journal of Agricultural Economics, 71*(1), 76-84.

Loomis, J.B. (1996). Measuring the economic benefits of removing dams and restoring the Elwha River: Results of a contingent valuation survey. *Water Resources Research, 32*(2), 441-447.

Loomis, J.B. & duVair, P.H. (1993). Evaluating the effect of alternative risk communication devices on willingness to pay: results from a dichotomous choice contingent valuation experiment. *Land Economics, 69*(3), 287-298.

Loomis, J., Gonzales-Caban, A., & Gregory, R. (1994). Do reminders of substitutes and budget constraints influence contingent valuation estimates? *Land Economics, 70*(4), 499-506.

Loomis, J. & King, M. (1994). Comparison of mail and telephone-mail contingent valuation surveys. *Journal of Environmental Management, 41*, 309-324.

Loomis, J.B., Sorg, C.F., & Donnelly, D.M. (1986). Evaluating regional demand models for estimating recreation use and economic benefits: a case study. *Water Resources Research, 22*(4), 431-438.

MacDonald, D.N., Murdoch, J.C., & White, H.L. (1987). Uncertain hazards, insurance, and consumer choice: evidence from housing markets. *Land Economics, 63*(4). 361-371.

MacDonald, H.F., Bergstrom, J.C., & Houston, J.E. (1998). A proposed methodology for measuring incremental environmental benefits from using constructed wetlands to control agricultural non-point-source pollution. *Journal of Environmental Management, 54*, 259-267.

Macmillan, D.C., Duff, E.I., & Elston, D.A. (2001). Modeling the non-market environmental costs and benefits of biodiversity projects using contingent valuation data. *Environmental and Resource Economics, 18*, 391-410.

Madariaga, B. & McConnell, K.E. (1987). Exploring existence value. *Water Resources Research, 23*(5), 936-942.

Mahan, B.L., Polasky, S., & Adams, R.M. (2000). Valuing urban wetlands: a property price approach. *Land Economics, 76*(1), 100-113.

Majid, I., Sinden, J.A., & Randall, A. (1983). Benefit evaluation of increments to existing systems of public facilities. *Land Economics, 59*(4), 377-392.

Mannesto, G. & Loomis, J.B. (1991). Evaluation of mail and in-person contingent value surveys: results of a study of recreational boaters. *Journal of Environmental Management, 32*, 177-190.

Mansfield, C. (1999). Despairing over disparities: explaining the difference between willingness to pay and willingness to accept. *Environmental and Resource Economics, 13*, 219-234.

Mark, J.H. (1980). A preference approach to measuring the impact of environmental externalities. *Land Economics, 56*(1), 103-116.

Maxwell, S. (1994). Valuation of rural environmental improvements using contingent valuation methodology: a case study of the Marston Vale Community Forest Project. *Journal of Environmental Management, 41*, 385-399.

McCluskey, J.J. & Rausser, G.C. (2001). Estimation of perceived risk and its effect on property values. *Land Economics, 77*(1), 42-55.

McConnell, K.E. (1977). Congestion and willingness to pay: a study of beach use. *Land Economics, 53*(2), 185-195.

McConnell, K.E. (1979). Values of marine recreational fishing: measurement and impact of measurement. *American Journal of Agricultural Economics, 61*(5), 921-925.

McFadden, D. (1994). Contingent valuation and social choice. *American Journal of Agricultural Economics, 76*(4), 689-708.

McLeod, D.M. & Bergland, O. (1999). Willingness-to-pay estimates using the double-bounded dichotomous-choice contingent valuation format: a test for validity and precision in a Baysian framework. *Land Economics, 75*(1), 115-125.

McMillan, M.L., Reid, B.G., & Gillen, D.W. (1980). An extension of the hedonic approach for estimating the value of quiet. *Land Economics, 56*(3), 315-328.

Messonnier, M.L., Bergstrom, J.C., Cornwell, C.M., Teasley, R.J., & Cordell, H.K. (2000). Survey response-related biases in contingent valuation: concepts, remedies, and empirical application to valuing aquatic plant management. *American Journal of Agricultural Economics, 83*(2), 438-450.

Michael, H.J., Boyle, K.J., &Bouchard, R. (2000). Does the measurement of environmental quality affect implicit prices estimated from hedonic models? *Land Economics, 76*(2), 283-298.

Milon, J.W., Gressel, J., & Mulkey, D. (1984). Hedonic amenity valuation and functional form specification. *Land Economics, 60*(4), 378-305.

Miyata, Y. & Abe, H. (1994). Measuring the effects of a flood control project: hedonic land price approach. *Journal of Environmental Management, 42*, 389-401.

Moore, J.L., Perin, D.E., & Maiden, B.G. (1979). Estimating the effect of water quality improvement on public swimming. *Water Resources Research, 15*(6), 1323-1328.

Morrison, M.D., Blamey, R.K., & Bennett, J.W. (2000). Minimising payment vehicle bias in contingent valuation studies. *Environmental and Resource Economics, 16*, 407-422.

Mullen, J.K. & Menz, F.C. (1985). The effect of acidification damages on the economic value of the Adirondack fishery to New York anglers. *American Journal of Agricultural Economics, 67*(1), 112-119.

Murray, C., Sohngen, B., & Pendleton, L. (2001). Valuing water quality advisories and beach amenities in the Great Lakes. *Water Resources Research, 37*(10), 2583-2590.

Nelson, A.C., Genereux, J., & Genereux, M. (1992). Price effects of landfills on house values. *Land Economics, 68*(4), 359-365.

Oster, S. (1977). Survey results on the benefits of water pollution abatement in the Merrimack River basin. *Water Resources Research, 13*(6), 882-884.

Ozuna, T., Jang, K.Y., & Stoll, J.R. (1993). Testing for misspecification in the referendum contingent valuation approach. *American Journal of Agricultural Economics, 75*(2), 332-338.

Palmquist, R.B. & Israngkura, A. (1999). Valuing air quality with hedonic and discrete choice models. *American Journal of Agricultural Economics, 81*(5), 1128-1133.

Palmquist, R.B., Roka, F.M., & Vukina, T. (1997). Hog operations, environmental effects, and residential property values. *Land Economics, 73*(1), 114-124.

Park, T. & Loomis, J. (1996). Joint estimation of contingent valuation survey responses. *Environmental and Resource Economics, 7*, 149-162.

Park, T.P., Loomis, J.B., & Creel, M. (1991). Confidence intervals for evaluating benefits estimates from dichotomous choice contingent valuation studies. *Land Economics, 67*(1), 64-73.

Patrick, R., Fletcher, J., Lovejoy, S., van Beek, W., Holloway, G., & Binkley, J. (1991). Estimating regional benefits of reducing targeted pollutants: an application to agricultural effects on water quality and the value of recreational fishing. *Journal of Environmental Management, 33*, 301-310.

Pattanayak, S.K. & Kramer, R.A. (2001). Pricing ecological services: willingness to pay for drought mitigation from watershed protection in eastern Indonesia. *Water Resources Research, 37*(3), 771-778.

Pendleton, L. & Mendelsohn, R. (2000). Estimating recreation preferences using hedonic travel cost and random utility models. *Environmental and Resource Economics, 17*, 89-108.

Piper, S. & Martin, W.E. (1997). Household willingness to pay for improved rural water supplies: a comparison of four sites. *Water Resources Research, 33*(9), 2153-2163.

Poe, G.L. (1998). Valuation of groundwater quality using contingent valuation - damage function approach. *Water Resources Research, 34*(12), 3627-3633.

Poe, G.L. & Bishop, R.C. (1999). Valuing the incremental benefits of groundwater protection when exposure levels are known. *Environmental and Resource Economics, 13*, 341-367.

Poe, G.L., Welsh, M.P., & Champ, P. (1997). Measuring the difference in mean willingness to pay when dichotomous choice contingent valuation responses are not independent. *Land Economics, 73*(2), 255-267.

Poon, L.C.L. (1978). Railway externalities and residential property prices. *Land Economics, 54*(2), 218-227.

Poor, J.P., Boyle, K.J., Taylor, L.O., & Bouchard, R. (2001). Objective versus subjective measures of water clarity in hedonic property value models. *Land Economics, 77*(4), 482-493.

Pope III, C.A. & Jones, J.W. (1990). Value of wilderness designation in Utah. *Journal of Environmental Management, 30*, 157-174.

Powell, J.R., Allee, D.J., & McClintock, C. (1994). Groundwater protection planning: impact of contingent valuation information. *American Journal of Agricultural Economics, 76*(5), 1068-1075.

Rahmatian, M. (1986). Extensions of the disaggregate bid experiment: variations in framing. *Journal of Environmental Management, 22*, 191-202.

Rahmatian, M. (1987). Component value analysis: air quality in the Grand Canyon National Park. *Journal of Environmental Management, 24,* 217-223.

Ready, R.C., Buzby, J.C., & Hu, D. (1996). Differences between continuous and discrete contingent value estimates. *Land Economics, 72*(3), 397-411.

Reaves, D.W., Kramer, R.A., & Holmes, T.P. (1999). Does question format matter? Valuing an endangered species. *Environmental and Resource Economics, 14*, 365-383.

Reiling, S.D., Boyle, K.J., Phillips, M.L., & Anderson, M.W. (1990). Temporal reliability of contingent values. *Land Economics, 66*(2), 128-134.

Reynolds, I.K. (1978). The relationship of land values to site characteristics - some implications for scenic quality management. *Journal of Environmental Management, 6*, 99-106.

Riddel, M. & Loomis, J. (1998). Joint estimation of multiple CVM scenarios under a double bounded questioning format. *Environmental and Resource Economics, 12*, 77-98.

Samples, K.C., Dixon, J.A., & Gowen, M.M. (1986). Information disclosure and endangered species valuation. *Land Economics, 62*(3), 306-312.

Sanders, L.D., Walsh, R.G., & Loomis, J.B. (1990). Toward empirical estimation of the total value of protecting rivers. *Water Resources Research, 26*(7), 1345-1357.

Scarpa, R. & Bateman, I. (2000). Efficiency gains afforded by improved bid design versus follow-up valuation questions in discrete-choice CV studies. *Land Economics, 76*(2), 299-311.

Shabman, L. & Bertelson, M.K. (1979). The use of development value estimates for coastal wetland permit decisions. *Land Economics, 55*(2), 213-222.

Shabman, L. & Stephenson, K. (1996). Searching for the correct benefit estimate: empirical evidence for an alternative perspective. *Land Economics, 72*(4), 433-449.

Shaw, W.D. & Shonkwiler, J.S. (2000). Brand choice and purchase frequency revisited: an application to recreation behavior. *American Journal of Agricultural Economics, 82*, 515-526.

Shechter, M., Reiser, B., & Zaitsev, N. (1998). Measuring passive use value. *Environmental and Resource Economics, 12*, 457-478.

Shulstad, R.N. & Stoevener, H.H. (1978). The effects of mercury contamination in pheasants on the value of pheasant hunting in Oregon. *Land Economics, 54*(1), 39-49.

Shultz, S.D. & Lindsay, B.E. (1990). The willingness to pay for groundwater protection. *Water Resources Research, 26*(9), 1869-1875.

Sieg, H., Smith, V.K., Banzhaf, H.S., & Walsh, R. (1999). The role of optimizing behavior in willingness-to-pay estimates for air quality. *American Journal of Agricultural Economics, 81*(5), 1112-1117.

Silberman, J., Gerlowski, D.A., & Williams, N.A. (1992). Estimating existence value for users and nonusers of New Jersey beaches. *Land Economics, 68*(2), 225-236.

Sinden, J.A. (1974). A utility approach to the valuation of recreational and aesthetic experiences. *American Journal of Agricultural Economics, 56*(1), 61-72.

Smith, V.K. (1996). Can contingent valuation distinguish economic values for different public goods? *Land Economics, 72*(2), 139-151.

Smith, V.K., Desvousges, W.H., & Fisher, A. (1986). A comparison of direct and indirect methods for estimating environmental benefits. *American Journal of Agricultural Economics, 68*(2), 280-290.

Smith, V.K., Zhang, X., & Palmquist, R.B. (1997). Marine debris, beach quality, and non-market values. *Environmental and Resource Economics, 10*, 223-247.

Soderqvist, T. (1998). Why give up money for the Baltic Sea? *Environmental and Resource Economics, 12*, 249-254.

Soguel, N.C. (1995). Costing the traffic barrier effect: a contingent valuation survey. *Environmental and Resource Economics, 6*, 301-308.

Stenger, A. & Willinger, M. (1998). Preservation value for groundwater quality in a large aquifer: a contingent-valuation study of the Alsatian aquifer. *Journal of Environmental Management, 53*, 177-193.

Stevens, T.H., DeCoteau, N.E., & Willis, C.E. (1997). Sensitivity of contingent valuation to alternative payment schedules. *Land Economics, 73*(1), 140-148.

Stevens, T.H., Echeverria, J., Glass, R.J., Hager, T., & More, T.A. (1991). Measuring the existence value of wildlife: what do CVM estimates really show? *Land Economics, 67*(4), 390-400.

Stevens, T.H., Glass, R., More, T., & Echeverria, J. (1991). Wildlife recovery: is benefit-cost analysis appropriate? *Journal of Environmental Management, 33*, 327-334.

Stevens, T.H., More, T.A., & Glass, R.J. (1994). Interpretation and temporal stability of CV bids for wildlife existence: a panel study. *Land Economics, 70*(3), 355-363.

Streever, W.J., Callaghan-Perry, M., Searless, A., Stevens, T., & Svoboda, P. (1998). Public attitudes and values for wetland conservation in New South Wales, Australia. *Journal of Environmental Management, 54*, 1-14.

Struyk, R.J. (1971). Flood risk and agricultural land values: a test. *Water Resources Research, 7*(4), 789-797.

Sutherland, R.J. & Walsh, R.G. (1985). Effect of distance on the preservation value of water quality. *Land Economics, 61*(3), 281-291.

Svento, R. (1993). Some notes on trichotomous choice discrete valuation. *Environmental and Resource Economics, 3*, 533-543.

Svento, R. (1999). On the asymmetry of the vagueness band in willingness to pay answers. *Environmental and Resource Economics, 14*, 151-163.

Taylor, L.O. & Smith, V.K. (2000). Environmental amenities as a source of market power. *Land Economics, 76*(4), 550-568.

Thibodeau, F.R. & Ostro, B.D. (1981). An economic analysis of wetland protection. *Journal of Environmental Management, 12*, 19-30.

Thomas, J.F. & Syme, G.J. (1988). Estimating residential price elasticity of demand for water: a contingent valuation approach. *Water Resources Research, 24*(11), 1847-1857.

Torell, L.A., Libbin, J.D., & Miller, M.D. (1990). The market value of water in the Ogallala aquifer. *Land Economics, 66*(2), 163-175.

Tyrvainen, L. (2001). Economic valuation of urban forest benefits in Finland. *Journal of Environmental Management, 62*, 75-92.

Vaughan, W.J. & Russell, C.S. (1982). Valuing a fishing day: an application of a systematic varying parameter model. *Land Economics, 58*(4), 450-461.

Walsh, R.G., Bjonback, R.D., Aiken, R.A., & Rosenthal, D.H. (1990). Estimating the public benefits of protecting forest quality. *Journal of Environmental Management, 30*, 175-189.

Walsh, R.G., Loomis, J.B., & Gillman, R.A. (1984). Valuing option, existence, and bequest demands for wilderness. *Land Economics, 60*(1), 14-29.

Walsh, R.G., Ward, F.A., & Olienyk, J.P. (1989). Recreational demand for trees in national forests. *Journal of Environmental Management, 28*, 255-268.

Ward, F.A., Roach, B.A., & Henderson, J.E. (1996). The economic value of water in recreation: evidence from the California drought. *Water Resources Research, 32*(4), 1075-1081.

Weicher, J.C. & Zerbst, R.H. (1973). The externalities of neighborhood parks: an empirical investigation. *Land Economics, 49*, 99-105.

White, P.C.L. & Lovett, J.C. (1999). Public preferences and willingness-to-pay for nature conservation in the North York Moors National Park, UK. *Journal of Environmental Management, 55*, 1-13.

Whitehead, J.C. (1995). Willingness to pay for quality improvements: comparative statics and interpretation of contingent valuation results. *Land Economics, 71*(2), 207-215.

Whitehead, J.C. & Blomquist, G.C. (1991). Measuring contingent values for wetlands: effects of information about related environmental goods. *Water Resources Research, 27*(10), 2523-2531.

Whitehead, J.C., Blomquist, G.C., Ready, R.C., & Huang, J. (1998). Construct validity of dichotomous and polychotomous choice contingent valuation questions. *Environmental and Resource Economics, 11*, 107-116.

Whitehead, J.C. & Hoban, T.J. (1999). Testing for temporal reliability in contingent valuation with time for changes in factors affecting demand. *Land Economics, 75*(3), 453-465.

Whitehead, J.C., Hoban, T.J., & Clifford, W.B. (1995). Measurement issues with iterated, continuous/interval contingent valuation data. *Journal of Environmental Management, 43*, 129-139.

Willis, K.G. & Garrod, G.D. (1993). Valuing landscape: a contingent valuation approach. *Journal of Environmental Management, 37*, 1-22.

Willis, K.G. & Garrod, G.D. (1999). Angling and recreation values of low-flow alleviation in rivers. *Journal of Environmental Management, 57*, 71-83.

Willis, K.G., Garrod, G.D., & Saunders, C.M. (1995). Benefits of environmentally sensitive area policy in England: a contingent valuation assessment. *Journal of Environmental Management, 44*, 105-125.

Xu, F., Mittelhammer, R.C., & Barkley, P.W. (1993). Measuring the contributions of site characteristics to the value of agricultural land. *Land Economics, 69*(4), 356-369.

Zabel, J.E. & Kiel, K.A. (2000). Estimating the demand for air quality in four U.S. cities. *Land Economics, 76*(2), 174-194.

REFERENCES

(2001). Army developing tool to measure impacts on natural resources. Defense Environment Alert via InsideEPA.com. 9.

Adamowicz, W.L., Bhardwaj, V., & Macnab, B. (1993). Experiments on the difference between willingness to pay and willingness to accept. *Land Economics, 69*(4), 416-427.

Alberini, A. (1995). Testing willingness-to-pay models of discrete choice contingent valuation survey data. *Land Economics, 71*(1), 83-95.

Aldred, J. & Jacobs, M. (2000). Citizens and wetlands: evaluating the Ely citizens' jury. *Ecological Economics, 34*, 217-232.

Arrow, K., Solow, R., Portney, P., Leamer, E., Radner, R., & Schuman, H. (1993). Report of the NOAA Panel on Contingent Valuation, National Oceanic and Atmospheric Administration: 64.

Atkinson, S.E., Crocker, T.D., & Shogren, J.F. (1992). Bayesian exchangeability, benefit transfer, and research efficiency. *Water Resources Research, 28*(3), 715-722.

Balistreri, E., McClelland, G., Poe, G., & Schulze, W. (2001). Can hypothetical questions reveal true values? A laboratory comparison of dichotomous choice and open-ended contingent values with auction values. *Environmental and Resource Economics, 18*, 275-292.

Barbier, E.B. (1994). Valuing environmental functions: tropical wetlands. *Land Economics, 70*(2), 155-173.

Bardach, E. (2000). *A Practical Guide for Policy Analysis*. New York: Chatham House Publishers.

Baron, J. & Spranca, M. (1997). Protected values. *Organizational Behavior and Human Decision Processes, 70*(1), 1-16.

Begg, C.B. (1994). Publication Bias. In H. Cooper & L.V. Hedges (Eds.), *The Handbook of Research Synthesis* (pp. 399-410). New York: Russel Sage Foundation.

Bengston, D.N., Fan, D.P., & Celarier, D.N. (1999). A new approach to monitoring the social environment for natural resource management and policy: the case of US national forest benefits and values. *Journal of Environmental Management, 56*, 181-193.

Bergstrom, J.C. (1990). Concepts and measures of the economic value of environmental quality: a review. *Journal of Environmental Management, 31*, 215-228.

Bishop, R., Champ, P.A., Brown, T.C., & McCollum, D.W. (1997). Measuring the non-use values: theory and application. In R.J. Kopp, W.W. Pommerehne, & N. Schwarz (Eds.) *Determining the Value of Non-Marketed Goods: Economic, Psychological, and Policy Relevant Aspects of Contingent Valuation Models*. Norell, MA: Kluwer Academic Publishers.

Bjornstad, D., Cummings, R., & Osborne, L. (1997). A learning design for reducing hypothetical bias in the contingent valuation method. *Environmental and Resource Economics, 10*, 207-221.

Blamey, R.K. (1998). Decisiveness, attitude expression and symbolic responses in contingent valuation surveys. *Journal of Economic Behavior & Organization, 34*(4), 577-601.

Blamey, R.K., Bennett, J.W., & Morrison, M.D. (1999). Yea-saying in contingent valuation studies. *Land Economics, 75*(1), 126-141.

Boyle, K.J. (1989). Commodity specification and the framing of contingent-valuation questions. *Land Economics, 65*(1), 57-63.

Boyle, K.J. & Bergstrom, J.C. (1992). Benefit transfer studies: myths, pragmatism, and idealism. *Water Resources Research, 28*(3), 657-663.

Boyle, K.J., Bergstrom, J.C., & Poe, G.L. (2001). Summary and conclusions. In J.C. Bergstrom, K.J. Boyle, & G.L. Poe (Eds.), *The Economic Value of Water Quality* (pp. 163-168). Northampton, MA: Edward Elgar.

Boyle, K.J., Johnson, F.R., McCollum, D.W., Desvousges, W.H., Dunford, R.W., & Hudson, S.P. (1996). Valuing public goods: discrete versus continuous contingent-valuation responses. *Land Economics, 72*(3), 381-396.

Boyle, K.J., MacDonald, H.F., Cheng, H., & McCollum, D.W. (1998). Bid design and yea saying in single-bounded, dichotomous-choice questions. *Land Economics, 74*(1), 49-64.

Boyle, K.J., Poe, G.L., & Bergstrom, J.C. (1994). What do we know about groundwater values? Preliminary implications from a meta-analysis of contingent valuation studies. *American Journal of Agricultural Economics, 76*(5), 1055-1061.

Bozeman, B. (2002). Public value failure: when efficient markets may not do. *Public Administration Review, 62*(2), 145-161.

Brady, G.L. (1983). The benefit/cost analytic framework: handling uncertainties in the Clean Air Act. *Journal of Environmental Management, 16*, 335-346.

Brookshire, D.S. & Neill, H.R. (1992). Benefit transfer: conceptual and empirical issues. *Water Resources Research, 28*(3), 651-655.

Brookshire, D.S., Thayer, M.A., Schulze, W.D., & d'Arge, R.C. (1982). Valuing public goods: a comparison of survey and hedonic approaches. *American Economic Review, 72*(1), 165-177.

Brouwer, R. (2000). Environmental value transfer: state of the art and future prospects. *Ecological Economics, 32*, 137-152.

Brower, D. (1995). Keynote address given on the 25th anniversary of the U.S. Environmental Protection Agency. San Francisco.

Brown, T.C., Champ, P.A., Bishop, R.C., & McCollum, D.W. (1996). Which response format reveals the truth about donations to a public good? *Land Economics, 72*(2), 152-166.

Brown, T.C. & Duffiend, J.W. (1995). Testing part-whole valuation effects in contingent valuation of instream flow protection. *Water Resources Research, 31*(9), 2341-2351.

Carson, R.T. (1997). Contingent valuation: theoretical advances and empirical tests since the NOAA panel. *American Journal of Agricultural Economics, 79*(5), 1501-1507.

Carson, R.T., Flores, N.E., Martin, K.M., & Wright, J.L. (1996). Contingent valuation and revealed preference methodologies: comparing the estimates for quasi-public goods. *Land Economics, 72*(1), 80-99.

Carson, R.T., Flores, N.E., & Meade, N.F. (2001). Contingent valuation: controversies and evidence. *Environmental and Resource Economics, 19*, 173-210.

Carson, R.T., Hanemann, W.M., Kopp, R.J., Krosnick, J.A., Mitchell, R.C., Presser, S., Ruud, P.A., Smith, V.K., Conaway, M., & Martin, K. (1997). Temporal reliability of estimates from contingent valuation. *Land Economics, 73*(2), 151-163.

Carson, R.T., Wright, J., Alberini, A., Carson, N., & Flores, N. (1994). A Bibliography of Contingent Valuation Studies and Papers. La Jolla, CA: Natural Resource Damage Assessment, Inc.: 142.

Champ, P.A. & Bishop, R.C. (2001). Donation payment mechanisms and contingent valuation: an empirical study of hypothetical bias. *Environmental and Resource Economics, 19*, 383-402.

Chattopadhyay, S. (1999). Estimating the demand for air quality: new evidence based on the Chicago housing market. *Land Economics, 75*(1), 22-38.

Chestnut, L.G., Ostro, B.D., & Vichit-Vadakan, N. (1997). Transferability of air pollution control health benefits estimates from the United States to developing countries: evidence from the Bangkok study. *American Journal of Agricultural Economics, 79*(5), 1630-1635.

Choe, K.C., Whittington, D., & Lauria, D.T. (1996). The economic benefits of surface water quality improvements in developing countries: a case study of Davao, Philippines. *Land Economics, 72*(4), 519-537.

Colwell, P.F. & Dilmore, G. (1999). Who was first? An examination of an early hedonic study. *Land Economics, 75*(4), 620-626.

Common, R., Reid, I., & Blamey, R. (1997). Do existence values for cost benefit analysis exist? *Environmental and Resource Economics, 9*, 225-238.

Cooper, H. & Hedges, L.V. (Eds.). (1994). *The Handbook of Research Synthesis*. New York: Russell Sage Foundation.

Costanza, R., d'Arge, R., de Groot, R., Farber, S., Grasso, M., Hannon, B., Limburg, K., Naeem, S., O'Neill, R.V., Paruelo, J., Raskin, R.G., Sutton, P., & van den Belt, M. (1997). The value of the world's ecosystem services and natural capital. *Nature, 387*(May), 253 - 260.

Cummings, R. & Taylor, L.O. (1998). Does realism matter in contingent valuation surveys? *Land Economics, 74*(2), 203-215.

Cummings, R.G. & Harrison, G.W. (1995). The measurement and decomposition of nonuse values: a critical review. *Environmental and Resource Economics, 5*, 225-247.

Dalecki, M.G., Whitehead, J.C., & Blomquist, G.C. (1993). Sample non-response bias and aggregate benefits in contingent valuation: and examination of early, late and non-respondents. *Journal of Environmental Management, 38*, 133-143.

David, E. L. (1971). Public perceptions of water quality. *Water Resources Research, 7*(3), 453-457.

Desvousges, W.H., Johnson, R.R., & Banzhaf, H.S. (1998). *Environmental Policy Analysis with Limited Information: Principles and Applications of the Transfer Method*. Northampton, MA: Edward Elgar.

Desvousges, W.H., Naughton, M.C., & Parsons, G.R. (1992). Benefit transfer: conceptual problems in estimating water quality benefits using existing studies. *Water Resources Research, 28*(3), 675-687.

Diamond, P.A. & Hausman, J.A. (1994). Contingent valuation: is some number better than no number? *Journal of Economic Perspectives, 8*(4), 45-64.

Downing, M. & Ozuna, T. (1996). Testing the reliability of the benefit function transfer approach. *Journal of Environmental Economics and Management, 30*, 316-322.

Ellis, R.D. (1998). *Just Results: Ethical Foundations for Policy Analysis*. Washington, DC: Georgetown University Press.

Ethier, R.G., Poe, G.L., Schulze, W.D., & Clark, J. (2000). A comparison of hypothetical phone and mail contingent valuation responses for green-pricing electricity programs. *Land Economics, 76*(1), 54-67.

Farber, S. (1988). The value of coastal wetlands for recreation: an application of travel cost and contingent valuation methodologies. *Journal of Environmental Management, 26*, 299-312.

Findley, R.W. & Farber, D.A. (1992). *Environmental Law in a Nutshell*. St. Paul: West Publishing Company.

Fleiss, J.L. (1994). Measures of Effect Size for Categorical Data. In H. Cooper & L.V. Hedges (Eds.), *The Handbook of Research Synthesis* (pp. 245-260). New York: Russell Sage Foundation.

Freeman, A.M., III (1993). *The Measurement of Environmental and Resource Values*. Washington, DC: Resources for the Future.

Freeman, A. M., III & Kopp, R.J. (1999). Assessing Damages from the Valdez Oil Spill. In W.E. Oats (Ed.), *The RFF Reader in Environmental and Resource Management* (pp. 51-54). Washington, DC: Resources for the Future.

Freeman, A. M., III & Portney, P.R. (1999). Economics Clarifies Choices about Managing Risk. In W.E. Oats (ed.), *The RFF Reader in Environmental and Resource Management* (pp. 15-20). Washington, DC: Resources for the Future.

Freeman, A. M., III. (1997). Economics, Incentives, and Environmental Regulation. In N. J. Vig and M.E. Kraft (Ed.), *Environmental Policy in the 1990s* (pp. 187-207). Washington, DC: CQ Press.

Gillroy, J.M. (1992). The ethical poverty of cost-benefit methods: autonomy, efficiency and public policy choice. *Policy Sciences, 25*, 83-103.

Gintis, H. (2000). Beyond homo economicus: evidence from experimental economics. *Ecological Economics, 35*, 311-322.

Giraud, K.L., Loomis, J.B., & Johnson, R.L. (1999). Internal and external scope in willingness-to-pay estimates for threatened and endangered wildlife. *Journal of Environmental Management, 56*, 221-229.

Goldar, B. & Misra, S. (2001). Valuation of environmental goods: correcting for bias in contingent valuation studies based on willingness-to-accept. *American Journal of Agricultural Economics, 83*(1), 150-156.

Goodin, R. (2003). A green theory of value. In J.J. Kassiola (Ed.), *Explorations in Environmental Political Theory: Thinking About What We Value* (pp. 128-148). Armonk, NY: M.E. Sharpe.

Goodland, R. & Ledec, G. (1994). Neoclassical Economics and Principles of Sustainable Development. In L.P. Pojman (Ed.), *Environmental Ethics: Readings in Theory and Application* (pp. 450-457). Sudbury, MA: Jones and Bartlett Publishers.

Gore, A. (1992). *Earth in the Balance: Ecology and the Human Spirit*. New York: Houghton Mifflin Company.

Gowdy, J.M. (1997). The value of biodiversity: markets, society, and ecosystems. *Land Economics, 73*(1), 25-41.

Green, C.H. & Tunstall, S.M. (1991). Is the economic evaluation of environmental resources possible? *Journal of Environmental Management, 33*, 123-141.

Gregory, R., Flynn, J., Johnson, S.M., Satterfield, T.A., Slovic, P., & Wagner, R. (1997). Decision-pathway surveys: a tool for resource managers. *Land Economics, 73*(2), 240-254.

Gregory, R., Lichtenstein, S., Brown, T.C., Peterson, G.L., & Slovic, P. (1995). How precise are monetary representations of environmental improvements? *Land Economics, 71*(4), 462-473.

Gregory, R., Lichtenstein, S., & Slovic, P. (1993). Valuing environmental resources: a constructive approach. *Journal of Risk and Uncertainty, 7*, 177-197.

Gregory, R., Mendelsohn, R., & More, T. (1989). Measuring the benefits of endangered species preservation: from research to policy. *Journal of Environmental Management, 29*, 399-407.

Hall, J.A., Tickle-Degnen, L., Rosenthal, R., & Mosteller, F. (1994). Hypotheses and problems in research synthesis. In H. Cooper & L.V. Hedges (Eds.), *The Handbook of Research Synthesis* (pp. 17-28). New York: Russell Sage Foundation.

Halvorsen, B. & Saelensminde, K. (1998). Differences between willingness-to-pay estimates from open-ended and discrete-choice contingent valuation methods: the effects of heteroscedasticity. *Land Economics, 74*(2), 262-282.

Hanley, N. (1992). Are there environmental limits to cost-benefit analysis? *Environmental and Resource Economics, 2*, 33-59.

Hanley, N., Spash, C., & Walker, L. (1995). Problems in valuing the benefits of biodiversity protection. *Environmental and Resource Economics, 5*, 249-272.

Hanley, N. & Spash, C.L. (1993). *Cost-Benefit Analysis and the Environment*. Northampton, MA: Edward Elgar Publishing.

Hardin, G. (1968). The Tragedy of the Commons. *Science, 162*(3858), 1243-48.

Hellerstein, D. (1995). Welfare estimation using aggregate and individual-observation models: a comparison using Monte Carlo techniques. *American Journal of Agricultural Economics, 77*, 620-630.

Hines, L.G. (1973). Precursors to benefit-cost analysis in early United States public investment projects. *Land Economics, 49*, 310-317.

Hite, D. (1998). Information and bargaining in markets for environmental quality. *Land Economics, 74*(3), 303-316.

Huang, J. & Smith, V.K. (1998). Monte Carlo benchmarks for discrete response valuation methods. *Land Economics, 74*(2), 186-202.

Hunt, M. (1997). *How Science Takes Stock: The Story of Meta-Analysis*. New York: Russell Sage Foundation.

Hunter, J.E. & Schmidt, F.L. (1990). *Methods of Meta-Analysis: Correcting Error and Bias in Research Findings*. Newbury Park: Sage Publications.

InsideEPA (2002). OMB's new rule cost-benefit methodology draws broad criticisms. *Inside EPA, 23*(6).

Jakobsson, K.M. & Dragun, A.K. (2001). The worth of a possum: valuing species with the contingent valuation method. *Environmental and Resource Economics, 19*, 211-227.

Johnston, R.J. & Swallow, S.K. (1999). Asymmetries in ordered strength of preference models: implications of focus shift for discrete-choice preference estimation. *Land Economics, 75*(2), 295-310.

Jordan, J.L. & Elnagheeb, A.H. (1994). Consequences of using different question formats in contingent valuation: a Monte Carlo study. *Land Economics, 70*(1), 97-110.

Jorgensen, B.S., Syme, G.J., Bishop, B.J., & Nancarrow, B.E. (1999). Protest responses in contingent valuation. *Environmental and Resource Economics, 14*, 131-150.

Kahneman, D. & Knetsch, J. (1992). Valuing public goods: the purchase of moral satisfaction. *Journal of Environmental Economics and Management, 22*, 57-70.

Karr, J.R. (2002). Biologic integrity and ecological health. In M.C. Newman & M.A. Unger (Eds.), *Fundamentals of Ecotoxicology*. Boca Raton, FL: CRC/Lewis Publishers.

Karr, J.R. & Rossano, E.M. (2001). Applying public health lessons to protect river health. *Ecology and Civil Engineering, 4*(1), 3-18.

Kask, S.B. & Shogren, J.F. (1994). Benefit transfer protocol for long-term health risk valuation: a case of surface water contamination. *Water Resources Research, 30*(10), 2813-2823.

Kassiola, J.J. (2003a). Introduction and overview: the nature of environmental political theory. In J.J. Kassiola (Ed.), *Explorations in Environmental Political Theory: Thinking About What We Value* (pp. 3-13). Armonk, NY: M.E. Sharpe.

Kassiola, J.J. (2003b). The "tragedy" of modernity: how environmental limits and the environmental crisis produce the need for postmodern values and institutions. In J.J. Kassiola (Ed.), *Explorations in Environmental Political Theory: Thinking About What We Value* (pp. 14-36). Armonk, NY: M.E. Sharpe.

Kealy, M.J., Dovidio, J.F., & Rockel, M.L. (1988). Accuracy in valuation is a matter of degree. *Land Economics, 64*(2), 158-171.

Keeble, J. (1999). *Out of the Channel: The Exxon Valdez Oil Spill in Prince William Sound*. Eastern Washington University Press.

Kellert, S.R. (1984). Assessing wildlife and environmental values in cost-benefit analysis. *Journal of Environmental Management, 18*, 355-363.

Kelman, S. (1998). Cost-benefit analysis: an ethical critique. In D. Van DeVer & C. Pierce (Eds.), *The Environmental Ethics and Policy Book* (pp. 316-322). Albany, NY: Wadsworth Publishing Company.

Kempton, W., Boster, J. & Hartley, J. (1995). *Environmental Values in American Culture*. Cambridge, MA: MIT Press.

Kerr, G.N. & Cullen, R. (1995). Public preferences and efficient allocation of a Possum-control budget. *Journal of Environmental Management, 43*, 1-15.

Kiel, K.A. (1995). Measuring the impact of the discovery and cleaning of identified hazardous waste sites on house values. *Land Economics, 71*(4), 428-435.

King, D. (1998). The dollar value of wetlands: trap set, bait taken, don't swallow. *National Wetlands Newsletter, 20*(4), 7-11.

Kirchoff, S., Colby, B.G., & LaFrance, J.T. (1997). Evaluating the performance of benefit transfer: an empirical inquiry. *Journal of Environmental Economics and Management, 33*, 75-93.

Kneese, A.V. (1999). The Faustian Bargain. In W.E. Oats (Ed.), *The RFF Reader in Environmental and Resource Management* (pp. 55-60). Washington, DC: Resources for the Future.

Knetsch, J.L. (1990). Environmental policy implications of disparities between willingness to pay and compensation demanded measures of values. *Journal of Environmental Management, 18*, 227-237.

Kolstad, C.D. & Guzman, R.M. (1999). Information and divergence between willingness to accept and willingness to pay. *Journal of Environmental Economics and Management, 38*, 215-233.

Kristrom, B. (1997). Practical problems in contingent valuation. In R.J. Kopp, W.W. Pommerehne, & N. Schwarz (Eds.), *Determining the Value of Non-Market Goods: Economic, Psychological, and Policy Relevant Aspects of Contingent Valuation Models*. Norell, MA: Kluwer Academic Publishers.

Leonard, H.B. & Zeckhauser, R.J. (1998). Cost-Benefit Analysis Defended. In D. Van DeVer & C. Pierce (Eds.), *The Environmental Ethics and Policy Book* (pp. 322-325). Albany, NY: Wadsworth Publishing Company.

Lipsey, M.W. (1994). Identifying Potential Interesting Variables and Analysis Opportunities. In H. Cooper & L.V. Hedges (Eds.), *The Handbook of Research Synthesis* (pp. 111-123). New York: Russell Sage Foundation.

Loehman, E.T., Park, S., & Boldt, D. (1994). Willingness to pay for gains and losses in visibility and health. *Land Economics, 70*(4), 478-498.

Loomis, J. (1986). Economic losses to recreational fisheries due to small-head hydro-power development: a case study of the Henry's Fork in Idaho. *Journal of Environmental Management, 22*, 85-94.

Loomis, J. (1987). The economic value of instream flow: methodology and benefit estimates for optimum flows. *Journal of Environmental Management, 24*, 169-179.

Loomis, J., Brown, T., Lucero, B., & Peterson, G. (1996). Improving validity experiments of contingent valuation methods: results of efforts to reduce the disparity of hypothetical and actual willingness to pay. *Land Economics, 72*(4), 450-461.

Loomis, J., Brown, T., Lucero, B., & Peterson, G. (1997). Evaluating the validity of the dichotomous choice question format in contingent valuation. *Environmental and Resource Economics, 10*, 109-123.

Loomis, J. & King, M. (1994). Comparison of mail and telephone-mail contingent valuation surveys. *Journal of Environmental Management, 41*, 309-324.

Loomis, J., Roach, R., Ward, R., & Ready, R. (1995). Testing transferability of recreation demand models across regions: a study of Corps of Engineers reservoirs. *Water Resources Research, 31*(3), 721-730.

Loomis, J.B. & Walsh, R.G. (1986). Assessing wildlife and environmental values in cost-benefit analysis: state of the art. *Journal of Environmental Management, 22*, 125-131.

Lyons, J., Navarro-Perez, S., Cochran, P.A., Santana, E., & Guzman-Arroyo, M. (1995). Index of biotic integrity based on fish assemblages for the conservation of streams and rivers in west-central Mexico. *Conservation Biology, 9*(3), 569-584.

Maguire, L. (2002). Remarks on "Thinking outside economics to value the environment." Annual Research Conference of the Association for Public Policy Analysis and Management. Dallas.

Mahan, B.L., Polasky, S., & Adams, R.M. (2000). Valuing urban wetlands: a property price approach. *Land Economics, 76*(1), 100-113.

Mannesto, G. & Loomis, J.B. (1991). Evaluation of mail and in-person contingent value surveys: results of a study of recreational boaters. *Journal of Environmental Management, 32*, 177-190.

Mansfield, C. (1999). Despairing over disparities: explaining the difference between willingness to pay and willingness to accept. *Environmental and Resource Economics, 13*, 219-234.

Markey, E. (2001). Morris K. Udall Arctic Wilderness Act of 2001. Washington, DC: U.S. House of Representatives.

Mattsson, L. & Li, C.Z. (1994). How do different forest management practices affect the non-timber value of forests? An economic analysis. *Journal of Environmental Management, 41*, 79-88.

McConnell, K. (1975). Some problems in estimating the demand for outdoor recreation. *American Journal of Agricultural Economics, 57*(2), 330-335.

McDaniels, T.L. & Roessler, C. (1998). Multiattribute elicitation of wilderness preservation benefits: a constructive approach. *Ecological Economics, 27*, 299-312.

McLaughlin, A. (2003). Industrialism and deep ecology. In J.J. Kassiola (Ed.), *Explorations in Environmental Political Theory: Thinking About What We Value* (pp. 104-127). Armonk, NY: M.E. Sharpe.

Messonnier, M.L., Bergstrom, J.C., Cornwell, C.M., Teasley, R.J., & Cordell, H.K. (2000). Survey response-related biases in contingent valuation: concepts, remedies, and empirical application to valuing aquatic plant management. *American Journal of Agricultural Economics, 83*(2), 438-450.

Michael, H.J., Boyle, K.J., & Bouchard, R. (2000). Does the measurement of environmental quality affect implicit prices estimated from hedonic models? *Land Economics, 76*(2), 283-298.

Milbrath, L.W. (2003). Envisioning a sustainable society. In J.J. Kassiola (Ed.), *Explorations in Environmental Political Theory: Thinking About What We Value* (pp. 37-51). Armonk, NY: M.E. Sharpe.

Mitchell, R.C. & Carson, R.T. (1989). *Using Surveys to Value Public Goods: The Contingent Valuation Method*. Washington, DC: Resources for the Future.

Moody, R.E. (1974). Inadequacy of the cost-benefit ratio as a measure of the public interest. *American Journal of Agricultural Economics, 56*(1), 188-191.

More, T.A., Averill, J.R., & Stevens, T.H. (1996). Values and economics in environmental management: a perspective and critique. *Journal of Environmental Management, 48*, 397-409.

Morgan, C. & Owens, N. (2001). Benefits of water quality policies: the Chesapeake Bay. *Ecological Economics, 39*, 271-284.

Morrison, M.D., Blamey, R.K., & Bennett, J.W. (2000). Minimizing payment vehicle bias in contingent valuation studies. *Environmental and Resource Economics, 16*, 407-422.

Munger, M.C. (2000). *Analyzing Policy: Choices, Conflicts, and Practices*. New York: W.W. Norton & Company.

Navrud, S. & Pruckner, G.J. (1997). Environmental valuation - to use or not to use? *Environmental and Resource Economics, 10*, 1-26.

Neill, H.R., Cummings, R.G., Ganderton, P.T., Harrison, G.W., & McGuckin, T. (1994). Hypothetical surveys and real economic commitments. *Land Economics, 70*(2), 145-154.

Norgaard, R.B. (1985). Environmental economics: an evolutionary critique and a plea for pluralism. *Journal of Environmental Economics and Management, 12*, 382-394.

Norgaard, R.B., Bode, C., & Values Reading Group (1998). Next, the value of God, and other reactions. *Ecological Economics, 25*, 37-39.

Norton, B.G. (1995). Evaluating ecosystem states: two competing paradigms. *Ecological Economics, 14*, 113-127.

Norton, B., Costanza, R., & Bishop, R.C. (1998). The evolution of preferences: why 'sovereign' preferences may not lead to sustainable policies and what to do about it. *Ecological Economics, 24*, 193-211.

Nunes, P.A.L.D. & van den Bergh, J.C.J.M. (2001). Economic valuation of biodiversity: sense or nonsense? *Ecological Economics, 39*, 203-222.

Nyborg, K. (2000). Project analysis as input to public debate: environmental valuation versus physical unit indicators. *Ecological Economics, 34*, 393-408.

Ofiara, D.D. & Seneca, J.J. (2001). *Economic Losses from Marine Pollution: A Handbook for Assessment*. Washington, DC: Island Press.

O'Hara, S.U. (1996). Discursive ethics in ecosystem valuation and environmental policy. *Ecological Economics, 16*, 95-107.

Paehlke, R. (2003). Environmentalism and Progressive Politics. In J.J. Kassiola (Ed.), *Explorations in Environmental Political Theory: Thinking About What We Value* (pp. 81-103). Armonk, NY: M.E. Sharpe.

Paris, D.C. & Reynold, J.F. (1983). *The Logic of Policy Inquiry*. New York: Longman.

Parsons, W. (1995). *Public Policy: An Introduction to the Theory and Practice of Policy Analysis*. Northampton, MA: Edward Elgar.

Patton, C.V. & Sawicki, D.S. (1993). *Basic Methods of Policy Analysis and Planning*. Englewood Cliffs, NJ: Prentice Hall.

Piper, S. & Martin, W.E. (2001). Evaluating the accuracy of the benefit transfer method: a rural water supply application in the USA. *Journal of Environmental Management, 63*, 223-235.

Prato, T. (1999). Multiple attribute decision analysis for ecosystem management. *Ecological Economics, 30*, 207-222.

Randall, A., Hoehn, J.P., & Brookshire, D.S. (1983). Contingent valuation surveys for evaluating environmental assets. *Natural Resources Journal, 23*(July), 635-648.

Ready, R.C., Buzby, J.C., & Hu, D. (1996). Differences between continuous and discrete contingent value estimates. *Land Economics, 72*(3), 397-411.

Ready, R.C., Navrud, S., & Dubourg, W.R. (2001). How do respondents with uncertain willingness to pay answer contingent valuation questions? *Land Economics, 77*(3), 315-326.

Reaves, D.W., Kramer, R.A., & Holmes, T.P. (1999). Does question format matter? Valuing an endangered species. *Environmental and Resource Economics, 14*, 365-383.

Rees, W.E. (1994). Sustainable Development: Economic Myths and Ecological Realities. In L.P. Pojman (Ed.), *Environmental Ethics: Readings in Theory and Application* (pp. 437-443). Sudbury, MA: Jones and Bartlett Publishers.

Rees, W.E. (1998). How should a parasite value its host? *Ecological Economics, 25*, 49-52.

Reisch, M. (2001). CRS Report for Congress: Superfund and Natural Resource Damages. Washington, DC: Congressional Research Service: 6.

Repetto, R. (1998). Earth in the Balance Sheet. In D. Van DeVer & C. Pierce (Eds.), *The Environmental Ethics and Policy Book* (pp. 327-334). Albany, NY: Wadsworth Publishing Company.

Rogers, P. (1995). Federal agency to quit putting price tag on oil spills. *San Jose Mercury News*. August 4.

Rosenbaum, W.A. (1998). *Environmental Politics and Policy*. Washington, DC: CQ Press.

Rosenberg, A. (1992). *Economics: Mathematical Politics or Science of Diminishing Returns?* Chicago: University of Chicago Press.

Rosenberger, R.S. & Loomis, J.B. (2000). Using meta-analysis for benefit transfer: in-sample convergent validity tests of an outdoor recreation database. *Water Resources Research, 36*(4), 1097-1107.

Rosenthal, D.H., Nelson, R.H., & Kopp, R.J. (1992). Insights: why existence values should not be used in cost-benefit analysis - comment/reply. *Journal of Policy Analysis and Management, 11*(1), 116-122.

Rosenthal, R. (1994). Parametric measures of effect size. In H. Cooper & L.V. Hedges (Eds.), *The Handbook of Research Synthesis* (pp. 231-244). New York: Russel Sage Foundation.

Rosenthal, R. & Rosnow, R. (1991). *Essentials of Behavioral Research: Methods and Data Analysis*. San Francisco: McGraw-Hill.

Sagoff, M. (1981). At the Shrine of Our Lady of Fatima, or why political questions are not all economic. *Arizona Law Review, 23*, 1283-1298.

Sagoff, M. (1994). Should preferences count? *Land Economics, 70*(2), 127-144.

Sagoff, M. (1997). The Economy of the Earth. In R.V. Percival & D.C. Alevizatos (Eds.), *Law and the Environment: A Multidisciplinary Reader* (pp. 49-53). Philadelphia: Temple University Press.

Sagoff, M. (1998). Aggregation and deliberation in valuing environmental public goods: a look beyond contingent pricing. *Ecological Economics, 24*(2-3), 213-230.

Sanders, L.D., Walsh, R.G., & McKean, J.R. (1991). Comparable estimates of the recreational value of rivers. *Water Resources Research, 27*(7), 1387-1394.

Satterfield, T., Slovic, P., & Gregory, R. (2000). Narrative valuation in a policy judgment context. *Ecological Economics, 34*, 315-331.

Scarpa, R. & Bateman, I. (2000). Efficiency gains afforded by improved bid design versus follow-up valuation questions in discrete-choice CV studies. *Land Economics, 76*(2), 299-311.

Schulze, W.D., d'Arge, R.C., & Brookshire, D.S. (1981). Valuing environmental commodities: some recent experiments. *Land Economics, 57*(2), 151-172.

Seip, K. & Strand, J. (1992). Willingness to pay for environmental goods in Norway: a contingent valuation study with real payment. *Environmental and Resource Economics, 2*, 91-106.

Shabman, L. & Stephenson, K. (1996). Searching for the correct benefit estimate: empirical evidence for an alternative perspective. *Land Economics, 72*(4), 433-449.

Shechter, M., Reiser, B., & Zaitsev, N. (1998). Measuring passive use value. *Environmental and Resource Economics, 12*, 457-478.

Shrestha, R.K. & Loomis, J.B. (2001). Testing a meta-analysis model for benefit transfer in international outdoor recreation. *Ecological Economics, 39*, 67-83.

Skalski, J.R. (1995). Statistical considerations in the design and analysis of environmental damage assessment studies. *Journal of Environmental Management, 43*, 67-85.

Smith, C.L. & Hogg, T.C. (1971). Benefits and beneficiaries: contrasting economic and cultural distinctions. *Water Resources Research, 7*(2), 254-263.

Smith, T.B., Bruford, M.W., & Wayne, R.K. (1996). The preservation of process: the missing element of conservation programs. In F.B. Samson & F.L. Knopf (Eds.), *Ecosystem Management: Selected Readings*. New York: Springer.

Smith, V.K. (1992). On separating defensible benefit transfers from "smoke and mirrors." *Water Resources Research, 28*(3), 685-694.

Smith, V.K. (1993). Nonmarket valuation of environmental resources: an interpretive appraisal. *Land Economics, 69*(1), 1-26.

Smith, V.K. (1996). Can contingent valuation distinguish economic values for different public goods? *Land Economics, 72*(2), 139-151.

Smith, V.K. & Huang, J.C. (1993). Hedonic models and air pollution: twenty-five years and counting. *Environmental and Resource Economics, 3*, 381-394.

Smith, V.K. & Huang, J.C. (1995). Can markets value air quality? A meta-analysis of hedonic property value models. *Journal of Political Economy, 103*(1), 209-227.

Smith, V.K. & Kaoru, Y. (1990). Signals or noise? Explaining the variation in recreation benefit estimates. *American Journal of Agricultural Economics, 72*(2), 419-433.

Smith, V.K & Pattanayak, S.K. (2002). Is meta-analysis a Noah's ark for non-market valuation? *Environmental and Resource Economics, 22*, 271-296.

Stevens, T.H., DeCoteau, N.E., & Willis, C.E. (1997). Sensitivity of contingent valuation to alternative payment schedules. *Land Economics, 73*(1), 140-148.

Stevens, T.H., Echeverria, J., Glass, R.J., Hager, T., & More, T.A. (1991). Measuring the existence value of wildlife: what do CVM estimates really show? *Land Economics, 67*(4), 390-400.

Stevens, T.H., Glass, R., More, T., & Echeverria, J. (1991). Wildlife recovery: is benefit-cost analysis appropriate? *Journal of Environmental Management, 33*, 327-334.

Stock, W.A. (1994). Systematic Coding for Research Synthesis. In H. Cooper & L.V. Hedges (Eds.), *The Handbook of Research Synthesis* (pp. 125-138). New York: Russell Sage Foundation.

Stone, D. (1997). *Policy Paradox: The Art of Political Decision Making*. New York: W.W. Norton & Company, Inc.

Tong, R. (1986). *Ethics in Policy Analysis*. Englewood Cliffs, NJ: Prentice Hall, Inc.

Trumbo, C.W., Markee, N.L., O'Keefe, G.J., & Park, E. (1999). Water policy, economics, and system analysis: antecedent precipitation as a methodological concern in attitude surveys on water conservation. *Water Resources Research, 35*(4), 1269-1275.

Unsworth, R.E. & Bishop, R.C. (1994). Assessing natural resource damages using environmental annuities. *Ecological Economics, 11*, 35-41.

U.S. House of Representatives Committee on Merchant Marine and Fisheries (1991). Negotiated Settlement from the Exxon Valdez Oil Spill. Washington, DC, U.S. House of Representatives: 72.

U.S. Senate Committee on Energy and Natural Resources (1999). Exxon Valdez oil spill: report together with additional views (to accompany S.711). Washington, DC, United States Senate: 14.

van den Bergh, J.C.J.M., Button, K.J., Nijkamp, P., & Pepping, G.C. (1997). *Meta Analysis in Environmental Economics*. Boston: Kluwer Academic Publishers.

Van Houtven, G. & Cropper, M.L. (1996). When is a life too costly to save? The evidence from U.S. environmental regulations. *Journal of Environmental Economics and Management, 30*, 348-368.

Van Houtven, G.L. & Cropper, M.L. (1999). When is a life too costly to save? The evidence from environmental Regulations. In W.E. Oats (Ed.), *The RFF Reader in Environmental and Resource Management* (pp. 43-49). Washington, DC: Resources for the Future.

Vatn, A. & Bromley, D.W. (1994). Choices without prices without apologies. *Journal of Environmental Economics and Management, 26*, 129-148.

Vogt, W.P. (1999). *Dictionary of Statistics & Methodology: A Nontechnical Guide for the Social Sciences*. Thousand Oaks, CA: Sage Publications.

Walsh, R.G., Johnson, D.M., & McKean, J.R. (1992). Benefit transfer of outdoor recreation demand studies, 1968-1988. *Water Resources Research, 28*(3), 707-713.

Walsh, R.G., Ward, F.A., & Olienyk, J.P. (1989). Recreational demand for trees in national forests. *Journal of Environmental Management, 28*, 255-268.

Weaver, R.D. (1996). Prosocial behavior: private contributions to agriculture's impact on the environment. *Land Economics, 72*(2), 231-247.

Weimer, D.L. & Vining, A.R. (1992). *Policy Analysis: Concepts and Practice*. Upper Saddle River, NJ: Prentice Hall.

White, H.D. (1994). Scientific Communication and Literature Retrieval. In H. Cooper & L.V. Hedges (Eds.), *The Handbook of Research Synthesis* (pp. 41-55), New York: Russell Sage Foundation.

White, L., Jr. (1967). The Historical Roots of Our Ecological Crisis. *Science, 155*(3767).

Whitehead, J.C. (1995). Willingness to pay for quality improvements: comparative statics and interpretation of contingent valuation results. *Land Economics, 71*(2), 207-215.

Whitehead, J.C. & Blomquist, G.C. (1991). Measuring contingent values for wetlands: effects of information about related environmental goods. *Water Resources Research, 27*(10), 2523-2531.

Whitehead, J.C., Blomquist, G.C., Ready, R.C., & Huang, J. (1998). Construct validity of dichotomous and polychotomous choice contingent valuation questions. *Environmental and Resource Economics, 11*, 107-116.

Whitehead, J.C. & Hoban, T.J. (1999). Testing for temporal reliability in contingent valuation with time for changes in factors affecting demand. *Land Economics, 75*(3), 453-465.

Willis, K.G. & Garrod, G.D. (1993). Valuing landscape: a contingent valuation approach. *Journal of Environmental Management, 37*, 1-22.

Woodward, R.T. & Wui, Y. (2001). The economic value of wetland services: a meta-analysis. *Ecological Economics, 37*, 257-270.

Young, D. (2001). Arctic Coastal Plain Domestic Energy Security Act of 2001. Washington, DC, U.S. House of Representatives.